MUTATION RATE DETERMINATION OF THE HUMAN MITOCHONDRIAL CONTROL REGION AND ITS IMPLICATIONS FOR FORENSIC IDENTITY TESTING

Joseph E. Warren, B.S., M.S.

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

May 2000

APPROVED:

Robert C. Benjamin, Co-Major Professor
Arthur Eisenberg, Co-Major Professor
Harrell Gill-King, Committee Member
Philip Hartman, Committee Member
John Planz, Committee Member
Earl G. Zimmerman, Chair of the Department of Biological Sciences
C. Neal Tate, Dean of the Robert B. Toulouse School of Graduate Studies

Warren, Joseph E., Mutation Rate Analysis of the Human Mitochondrial D-loop and its Implications for Forensic Identity Testing. Doctor of Philosophy (Molecular Biology), May 2000, 106 pp., 13 tables, 13 figures, reference list, 75 titles.

To further facilitate mitochondrial DNA (mtDNA) sequence analysis for human identity testing, a better understanding of its mutation rate is needed. Prior to the middle 1990's the mutation rate applied to a forensic or evolutionary analysis was determined by phylogenetic means, This method involved calculating genetic distances as determined by amino acid or DNA sequence variability within or between species. The mutation rate as determined by this method ranged from 0.025-0.26 nucleotide substitutions/ site/ myr (million years). With the recent advent of mtDNA analysis as a tool in human identity testing an increased number of observations have recently come to light calling into question the mutation rate derived from the phylogenetic method. The mutation rate as observed from forensic analysis appears to be much higher than that calculated phylogenetically. This is an area that needs to be resolved in human identity testing. Mutations that occur within a maternal lineage can lead to a possible false exclusion of an individual as belonging to that lineage. A greater understanding of the actual rate of mutation within a given maternal lineage can assist in determining criteria for including or excluding individuals as belonging to that lineage. The method used to assess the mutation rate in this study was to compare mtDNA sequences derived from the HVI and HVII regions of the D-loop from several different maternal lineages. The sequence information was derived from five unrelated families consisting of thirty-five individuals. One intergenerational mutational event was found. This derives to approximately 1.9

ACKNOWLEDGMENTS

I would like to thank the following individuals for their assistance and cooperation in various phases of this project: JEB Stewart, Mitchell Holland, Pamela Pogue, Amy Smuts, Gina Pineda, Patrick Cooke, Kelley Wakefield, Catherine Deshotel, Amy Goldston, George Schiro, Bill Watson, Joe DiZinno, H. Gill-King, Robert Benjamin, Philip Hartman, Arthur Eisenberg, Gerard A. O'Donovan, John Planz, Patricia Gibson, Tom Parsons, Nicholina Warren, and Joseph Warren, Sr.

I would also like to thank GeneScreen Corporation (Dallas, TX), BioSynthesis Corporation (Lewisville, TX), and ReliaGene Technologies, Incorporated (New Orleans, LA) for use of their facilities and expertise.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The Human Mitochondrial Genome

Human mitochondria possess multiple copies of a circular DNA moleculewith16,569 base pairs. This circular DNA replicates independently of the cell's nuclear DNA (Anderson *et al*., 1981). Utilizing cesium chloride gradient centrifugation, it was discovered that mitochondrial DNA was made up of a heavy (H) strand rich in guanine, and a light (L) strand rich in cytosine (Budowle *et al*., 1990).

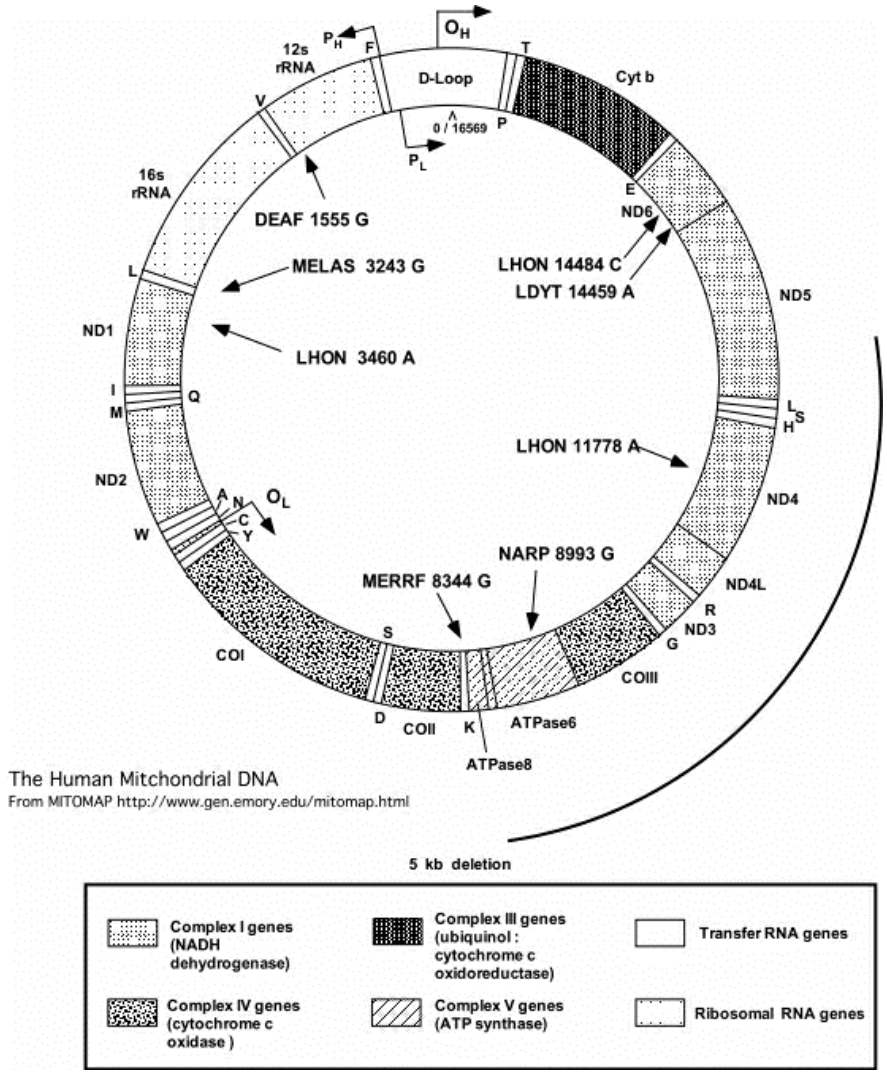
The human mitochondrial genome encodes 37 genes. Thirteen genes encode polypeptides. These include NADH dehydrogenase subunits ND1-ND6 and ND4L, cytochrome b, cytochrome c, oxidase subunits CAI-III and HEP synthetase subunits ATP8 and ATP6. Also encoded are the 12S and 16S rRNAs and 22 tRNAs (Anderson *et al*., 1981; **Fig**. **1**).

Replication

Mitochondria contain their own DNA polymerase (mtDNA polymerase) responsible for replication of the mitochondrial genome (Budowle *et al*., 1990). As mentioned previously, the two strands of the human mtDNA molecule are known as the H and L strands, and each strand has its own origin of replication. These have been designated O_H for the heavy strand and O_L for the light strand origins of replication. The O_H strand is located in the displacement loop (D-loop) region. The D-loop is a 680 non-coding nucleotide sequence flanked by phenylalanine tRNA and proline tRNA genes.

The conventional numbering scheme uses the O_H as the starting point for the numeric base designations.

Fig. 1. Human Mitochondrial DNA



Replication of the mtDNA molecule initiates at O_H, making H the leading strand. The mtDNA D-loop is somewhat unique because of its triplex nature. A short piece of daughter H-strand remains associated with the L-strand at all times. When H-strand replication is two-thirds complete, O_L is exposed as single-stranded DNA by the H-strand replication fork and this initiates the lagging strand replication.

Transcription of mtDNA

The D-loop region is also the control site for transcription of human mtDNA, as well as the site for initiation of its replication. The L and H strands have their own unique promoters; namely the L-strand promoter (LSP) and the H-strand promoter (HSP). The LSP is the promoter for NADH6 and eight tRNA genes. Transcription initiated at the LSP also serves as the primer (described above) for the leading strand mtDNA replication. The transition from DNA synthesis to RNA transcription in the mtDNA genome occurs over a series of short conserved sequences called conserved sequence blocks (CSB I-III) (Larson & Clayton, 1995). Mitochondrial RNA polymerase, mitochondrial transcription factor A (h-mtTFA), and perhaps some additional factors, are needed for mtDNA transcription. The h-mtTFA binds at a region upstream of both the HSP and LSP. H-mtTFA can unwind and bend DNA, thus suggesting a manner in which the initiation of transcription could occur (Fisher *et al*., 1992; Shadel & Clayton, 1993).

Translation in Mitochondria

The most interesting feature of translation in mitochondria is that it does not utilize the universal genetic code. The termination codon UGA is read as the amino acid tryptophan in human mtRNAs. The codon AUA, which normally codes for isoleucine, is read as methionine. AUG and AUU serve as initiation codons for mtDNA translation.

There are only 22 tRNAs in the mitochondrial genetic code translation system. Each tRNA can read two or four synonymous codons. There is also an absence of AGA and AGG arginine codons since these are also stop codons (along with the usual nuclear DNA stop codons UAG and UAA) in mitochondrial mRNAs (Barrell *et al*., 1980; **Table 1**).

Table 1. Codon Variations in Mitochondrial DNA and Nuclear DNA

Codon Variations	**Mitochondrial DNA (yeast)**	**Nuclear DNA (yeast)**
Stop Codons	AGA, AGG, UAG, UAA	UAA, UAG, UGA
Start Codons	AUA, AUG	AUG
Amino Acids	UGA- Tryptophan	AGA, AGG- Arginine AUA- Isoleucine

Evolution of Mitochondria and mtDNAs

The most widely accepted scenario for the origin of mitochondria is the endosymbiont theory (Gray, 1989). Briefly stated, ancestors of mitochondria were at one time free- living ,bacteria-like organisms that were incorporated into a heteromorphic prokaryote. Based on comparative rRNA sequence analysis, their rRNA genes appear to closely resemble those found in some eubacterial genomes. This suggests that an early eubacterium was engulfed by a protoeukaryotic cell (Gray,1989). Eventually, the cells developed mechanisms to utilize the ATP of the bacteria, and the rich cellular environment enabled the bacteria to streamline its many biosynthetic pathways. The relationship also offered protection to the protomitochondrion and an energy source for the host bacteria and an overall selective advantage to both genomes. Over time, genes

were most likely exchanged between the nuclear and mitochondrial genomes with the net result being that nuclear DNA became more complex and the mtDNA became simple.

In accordance with the above hypothesis and based upon extensive genetic analysis, mitochondria are now divided into two evolutionary categories: ancestral and derived. The ancestral mitochondrial genomes have retained clear-cut vestiges of their eubacterial ancestry. These include:

1. The presence of extra genes (especially ribosomal and protein);
2. Genes that encode eubacterial-like rRNAs (23S, 16S, and 5S);
3. A complete, or almost complete, set of tRNA genes;
4. Tight packing of encoding genes with a few or no introns present;
5. Eubacterial-like gene clusters;
6. Standard genetic code
7. Large size

The genome of the Rickettsia is the prototypical example of an ancestral protomitochondria. Most animal,fungal and green algae mitochondria fall into the derived category. For review see Gray *et al*., 1999. Characteristics of the derived mitochondrial genomes depart dramatically from those of the ancestral mode. The derived mitochondrial genomes exhibit little or no evidence of retained primitive traits. Their characteristics include:

1. Extensive gene loss (both in proteins and tRNA genes);
2. Divergence of ribosomal DNA and rRNA structure. This is usually seen as truncation and fragmentation of rRNA sequences;
3. Accelerated rate of sequence divergence

4. Divergence from the universal genetic code.

Mitochondrial Genetics

The non-Mendelian, exclusively maternal, mode of inheritance is the most notable feature of human mtDNA (Case & Wallace, 1981; Giles *et al*., 1980). Although some paternal mtDNA from spermatozoa appear to enter the ovum at conception, this DNA seems to be eliminated soon after fertilization (Ankel-Simmon & Cummings 1996; Manfredi *et al*., 1997), however the mechanism for this is not known. Mature egg cells contain 10^5 mitochondria, sperm cells approximately 50 (Ferris *et al*., 1981a). The egg mitochondria are likely to overwhelm the sperm mitochondria in number, diluting any possible contributions.. Other speculations are that the egg cells contain nucleases that destroy sperm mtDNA or that the replication of sperm mtDNA is somehow impaired by the egg cell.

Mitochondrial DNA appears to undergo little or no intermolecular recombination (Hayashi *et al*., 1985) despite a recent study demonstrating that extracts from human mitochondria contain enzymes that can catalyze homologous recombination of plasmid DNA (Thayagarajan *et al*., 1996). These findings introduce a new paradox into our understanding of mtDNA: Why do human mitochondria possess a recombinase when repeated and recent efforts fail to find evidence of recombination in human mtDNA (Bidocki *et al*., 1997)?

Two other components of mtDNA genetics, homoplasmy (one type of mtDNA per cell per individual) and nucleotide substitution rates of mtDNA, will be discussed in a later section.

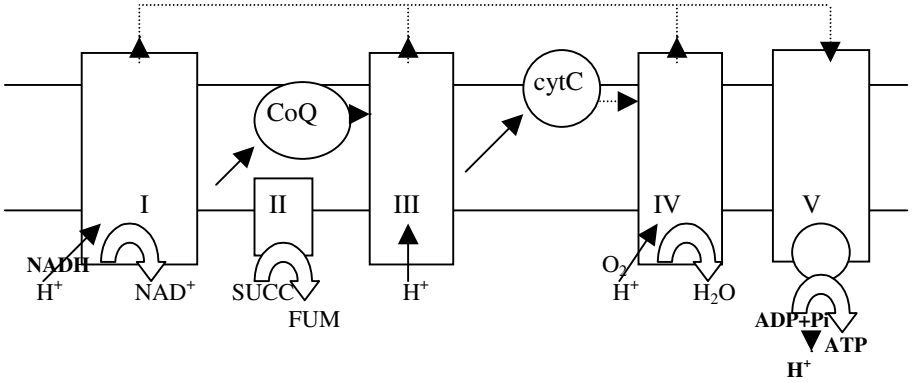
Medical Implications of Mitochondrial DNA Mutations

and the Pathologies Caused by These

The primary function of mitochondria is to produce ATP. To briefly summarize, the respiratory chain is located on the inner mitochondrial membrane and consists of approximately 90 different polypeptides divided into five enzyme complexes. Thirteen of these subunits are encoded by the mitochondrial genome.

The process by which ATP is produced is called oxidative phosphorylation. Protons are pumped out of the mitochondrial matrix by enzyme complexes I, III and IV. A proton gradient forms across the inner mitochondrial membrane. The protons then reenter the matrix through complex V (ATP synthetase). The proton gradient produces the energy needed for ATP synthesis (Larson & Clayton, 1995; **Fig. 2**).

Fig. 2. Schematic Illustration of Structure and Function of Respiratory Chain.



Note: The respiratory chain consists of five different enzyme complexes (complexes I-V), coenzyme Q (CoQ) and cytochrome C (cytC). NADH and Succinate are oxidized by complexes I and II. The electrons are transferred to coenzyme Q, complex III. Cytochrome c, complex IV, and finally to molecular oxygen which is reduced to water. Protons are pumped out of the mitochondrial matrix by complexes I, III, and IV and a proton gradient is formed across the inner membrane of the mitochondrion. The protons reenter the matrix through complex V (ATP synthase) and the energy of the proton gradient is used to synthesize ATP.

Recently, a number of diseases have been associated with changes in the sequence of the mitochondrial genome and the organelle's capacity to produce ATP.

These diseases are by and large neuromuscular afflictions. They are, as would be expected, commonly associated with those tissues that require a high amount of energy. Some examples are diseases of the central nervous system, such as Lebers Hereditary Optic Neuropathy (LHON), afflictions of skeletal and heart muscle, as well as problems associated with the kidney and liver. Specific examples include myoclonic epilepsy, ragged red fiber disease, mitochondrial encephalomyopathy and Kearns-Sayre Syndrome (KSS). These diseases, which are caused by mutations of the mtDNA, are classified into four main categories based upon the specific type of mutation: (Wallace,1989; Larson & Clayton, 1995)

1. missense mutations
2. biogenesis mutations affecting protein synthesis
3. insertion / deletion mutations
4. copy number mutations

A brief description of each of these categories with some examples of diseases found in them follows (**Table 2**).

Table 2. Genetic-based human diseases involving mtDNA mutations.

Type of Mutation	**Name of Disease**
Missense Mutations	LHON (Leber's Hereditary Optic Neuropathy) NARP (Neurogenic Muscle Weakness Ataxia and Retinosa Pigmentosa)
Biogenesis Mutations	MERRF (Myoclonic Epilepsy and Ragged Red Fiber Disease) MELAS (Mitochondrial Encephalomyopathy Lactic Acidosis and Stroke-like Symptons) MMC (Maternally Inherited Myopathy and Cardiomyopathy)
Insertion/Deletion Mutations	Ocular Myopathy Pearson's Syndrome KSS (Kearns-Sayre Syndrome)
Copy Number Mutations	
Autosomal Dominant Defects	

Missense Mutations

Missense mutations are point mutations that introduce a base pair change resulting in the coding of a new amino acid. There are two main phenotypes associated with missense mutations of the mitochondrial genome: LHON and NARP.

LHON is characterized by adult-onset blindness, usually at around 27 years of age, and is sometimes associated with a variety of cardiac effects. Male victims typically outnumber females 4 to 1. Fifty percent of LHON mutations are caused by the change of arginine to a histidine at amino acid position 340 in the NADH dehydrogenase subunit 4 (ND4) gene. (Wallace, 1989) This syndrome is readily detected because it changes a restriction enzyme cutting site from *Sfan*I to *Mae*I.

The remaining LHON cases are caused by eight missense mutations that have localized in three complex I genes and the cytochrome b gene of complex III. These mutations are judged to exist based on three criteria:

1. The presence of mutations at a higher frequency in patients than in controls
2. An alteration of evolutionarily conserved amino acids
3. Heteroplasmy (the presence of mtDNAs with the two different sequences in the same tissue)

These mutations can either cause LHON by themselves or can act synergistically depending on the particular mutation (Wallace, 1989). The reason for this effect might be that by identifying LHON mutations in both complex I and complex III genes, and also discovering that some of these mutations are cumulative. It appears that blindness might be more of a function of the total extent of inhibition of the electron transport system than the loss of any one particular enzyme.

The sex bias of LHON might be due to several factors. Both mitochondrial genes and nuclear DNA contribute to the proteins of the electron transport system. Therefore, a deleterious X-linked gene might be the cause of the higher number of males afflicted with LHON. LHON might also be a sex-limited trait if males are found to have a higher optic nerve metabolic rate than females. Environmental factors, such as exposure to electron transport chain inhibitors like cyanide and carbon monoxide, might also contribute to LHON formation. This suggests that the total cause for LHON in any particular patient might be quite complex (Wallace, 1989; Larson & Clayton, 1995).

NARP, which has a series of symptoms including neurogenic muscle weakness, ataxia seizures, retinosis pigmentosa, sensory neuropathy and developmental delay, has been associated with a missense mutation in the ATPase 6 gene (Wallace, 1989). This is caused by a thymine to guanine transversion mutation which changes a leucine to arginine at position number 156. Individuals demonstrate heteroplasmic maternal lineages, and the severity of the symptoms depends on the percentage of wild type and mutant mitochondrial genes (Wallace, 1989).

Biogenesis Mutations

Most of the biogenesis mutations are found in the biosynthetic genes that code for tRNAs (Wallace, 1989). The phenotypic symptoms are more general than those of the missense mutations. Biogenesis mutations are all associated with mitochondrial myopathy, which includes ragged red muscle cell fibers and abnormally formed mitochondria (Wallace, 1989) The more severe victims show nervous system, muscle, heart, and kidney involvement. These tRNA mutation diseases include myoclonic epilepsy and ragged red fiber disease (MERRF), mitochondrial encephalomyopathy,

lactic acidosis, and stroke-like symptoms (MELAS), and maternally inherited myopathy and cardiomyopathy (MMC) (Wallace, 1989).

MERRF is maternally inherited with affected individuals showing myoclonic epilepsy characterized by periodic jerking and mitochondrial myopathy. Hearing loss, dementia, respiratory failure, and renal dysfunction can also be seen. MERRF pedigrees are associated with pleiotropic defects of oxidative phosphorylation affecting complexes I and IV. The severity of the enzyme defect is directly proportional to the severity of the symptoms. This defect results in a reduction of protein synthesis in the mitochondria (Wallace, 1989). Most individuals are initially phenotypically normal and begin to show symptoms that get progressively worse as they age, due to the fact that as a person ages, the oxidation phosphorylation capacity of the tissues declines. The initial capacity is dependent on the original percentage of mutant mtDNA present at birth (Wallace, 1989; Larson & Clayton ,1995).

MELAS patients have reversible stroke-like symptoms with mitochondrial myopathy. The patients are diagnosed via CAT scanning or MR imaging along with examination of the other familial maternal relatives. MELAS is associated with defects in complex I genes. The main defects alter the dihydrouridine loop in tRNAs, which, in turn, deactivates transcriptional termination within that particular tRNA gene and all the rRNA genes downstream of it (Wallace, 1989). The result is reduced translation and the alternation of the ratio of rRNA and mRNA transcripts.

MMC is symptomatic for mitochondrial myopathy combined with hypertrophic cardiomyopathy. The cause is a combination of complex I and IV oxidative phosphorylation defects which result from the same tRNA mutations as does MELAS

(Wallace, 1989; Larson & Clayton, 1995). However, this particular mutation occurs outside the stem of the anti-codon loop and outside the transcriptional terminator sequence. Like many other mitochondrial diseases, MMC is heteroplasmic, and the severity of the symptoms is dependent on the proportion of mutant to wild type mtDNA (Wallace, 1989).

Insertion-Deletion Mutations

Two diseases of mitochondrial inheritance, which are the result of insertion-deletion mutations, are ocular myopathy and Pearsons Syndrome. A large range of symptoms is found in ocular myopathy,and consists of optholomolegia, ptosis, and mitochondrial myopathy to retinitis pigmentosa, lactic acidosis, hearing loss, ataxia, and heart conductive defects. The latter symptoms are grouped together and called Kearns-Sayre Syndrome (KSS) (Wallace, 1989).

Pearsons marrow/pancreas Syndrome is usually a childhood affliction and is characterized by loss of all blood cells, pancreatic fibrosis, and splenic atrophy. Those who survive childhood usually subsequently express a KSS phenotype. There have been more than 100 single mtDNA deletions characterized so far between the two diseases (Wallace, 1989). The size and position of the deletions vary among the patients. What all these deletions seem to have in common is that they have been confined to two sectors delineated by the H-strand and L-strand origins of replication. Some proposed molecular mechanisms for their origins include; slipped mispairing between distant repeats during mtDNA replication, brakage and ligation of replication intermediates, breakage and ligation of simple mature molecules and recombination between distant repeats (Larson & Clayton, 1995).

While studying tissues of these patients, it has been recorded that these deletions vary widely among their tissue distribution within individuals, which could account for the variety of symptoms that these diseases show. This also suggests that most of these mutations are new and occur during development because tissue lineages that are derived from cells prior to the deletion should have normal DNA.

Copy Numbered Mutations

Symptoms resulting from low copy number of mitochondria in the cell include lethal infantile respiratory failure, lactic acidosis, and muscle, liver, and kidney failures. Some victims of this disease have had mitochondrial levels of only 2% to 17% of the normal amount. It is interesting to note that AIDS patients treated with Zidovudine, a drug that blocks viral and mitochondrial DNA replication, produces a phenocopy of the mitochondrial copy number disease (Larson & Clayton, 1995).

Autosomal Dominant Effects

Certain mitochondrial deletions that cause ocular myopathy seem to be associated with autosomal dominant mutations occurring in the nuclear genes that code for proteins governing mitochondrial replication. Affected individuals usually show multiple deletions, as opposed to the normally seen single deletion (Larson & Clayton, 1995; Wallace, 1989; Wallace, 1999).

The Use of Human Mitochondrial DNA Analyses in Evolutionary Studies

Human beings have long pondered their origins. This is reflected in the ancient Babylonian epic of Gilgamesh and the Judeo-Christian account of creation in the Bibles Book of Genesis. Many other cultures and religions worldwide have their accounts of human origins.

Until recent times, the most direct and scientific account of human origins came from the fossil record. Skeletal remains, ancient tools, cave paintings etc., have proven to be invaluable in establishing the evolution of archaic humans in Africa, as well as the appearance and culture of modern humans (Fagan, 1990). Despite its impressive accomplishments, the fossil record remains spotty and incomplete.

When it became possible to measure genetic variation in humans, biologists began to use genetic data to study the affinities and origins of humans. Events that have effected the biology of human populations such as migration, bottlenecks and expansions all have leave imprints in the form of altered gene frequencies on the collective human genome (Jorde, 1998). Heredity transmits these imprints down to succeeding generations and thus the modern human genome can be thought of as a record of our evolutionary past.

Initially, blood groups, allozyme genetic markers and other protein polymorphisms where used in studying evolutionary genetic variation in human populations (Avise, 1994). Though providing important these markers were few in number, uniform in many populations and effected by natural selection (Jorde, 1998).

The advent of molecular techniques has led to the use of thousands of new genetic systems that can be quickly and easily assayed to provide information on different aspects of human evolution. Among these molecular based markers is mtDNA.

The mitochondrial genome offers a very different perspective on human evolution than other systems. Because mtDNA is inherited only through the maternal cytoplasm variation in mtDNA provides a record of the maternal lineage of our species. Since recombination does not happen in mtDNA therefore, the difference between any two

mitochondrial DNA sequences represent only the mutations that have occurred since each sequence was derived from a common ancestor. The non-coding control region of the mitochondrial genome has a high mutation rate compared to nuclear DNA, one estimate being 2-3 X10-7 (Horai, *et al*., 1995).

The origin of anatomically modern humans has been one of the more contentious and fascinating issues in evolutionary studies. The controversy revolves around two competing hypothesis. They are the "Multi-Regional Hypothesis" and the " African Replacement Hypothesis". The multi-regional hypothesis states that modern *Homo sapiens* evolved from more ancient forms over millions of years in several different old world locations. The genetic homogeneity observed in humans was maintained by gene flow and natural selection. The African replacement hypothesis states that modern humans arose in Africa 100,000 to 200,000 years ago to replace *Homo erectus* species completely (**Fig. 3**; Jorde, 1998). The debate over these differing schools of thought raise several issues that have been addressed by mtDNA analysis.

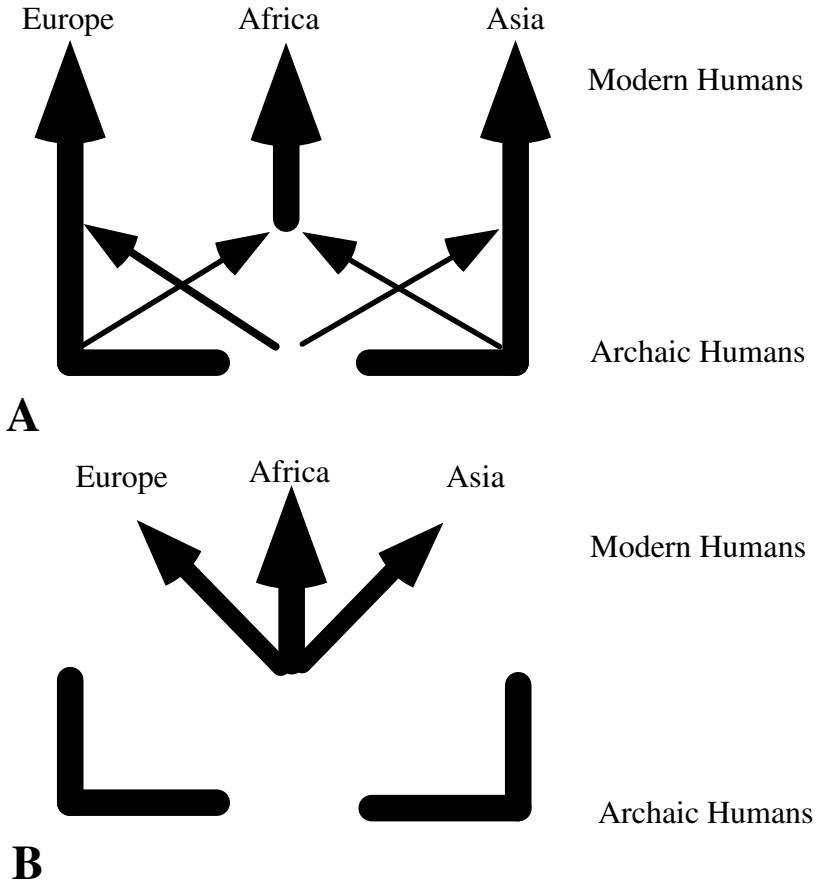


Fig. (**A**): A depiction of te multiregional hypothesis in which archaic humans evolve in situ into modern humans in Africa, Asia, and Europe. Gene flow, shown by black arrows, maintains genetic similarities between populations. (**B**): A depiction of the African replacement hypothesis in which modern humans evolve from archaic forms only in Africa. Archaic humans living in Asia and Europe are replaced by modern humans migrating out of Africa.

One of those issues is why do humans exhibit a relatively high level of genetic homogeneity when compared to other primate species ? Chimpanzees for example, exhibit variation in 1 out of every 5 nucleotides in the mitochondrial control region. Humans exhibit variation in 1 out of 17 nucleotides (Morin, *et al*., 1994). Studies on nuclear genes have also demonstrated this lower level of homogeneity in chimpanzees (Crouau-Roy, *et al*., 1996). The reduction of genetic variability in humans is more

compatible with the African replacement model that suggests that modern humans originated and/or expanded from a small population fairly recently (ie. the greater the genetic variation within a species, the longer it has existed).

The two most likely explanations for this low genetic(nucleotide) diversity are that either humans have recently undergone a major bottleneck in population size, or the effective population size of the human species has been small for a long time(Jorde, 1998).

Studies involving mtDNA sequences (Takahata, 1993), Y-chromosome data (Hammer, 1995) and nuclear DNA sequences (Takahata, 1995) all indicate that this bottleneck did occur and that the effective population size was approximately 10,000 individuals. This genetic data then appears to support a small effective population size consistent with a moderate bottleneck and a low level of genetic diversity.

Another useful parameter is to exam not only the genetic diversity of the entire human species, but also the diversity within and between major continental populations (also known as “races”) (Jorde, 1998). The way these intra and inter diversity is typically measured is using Wrights Fst statistic (Jorde, 1998). The Fst statistic for human mtDNA is 0.14 or 14% of the variation in mtDNA occurs between populations. The remaining 86% is found within populations (Jorde, 1998).

This Fst is a little larger then those reported for nuclear markers(0.04 to 0.11) (Jorde, 1995; Jorde, 1997). This result reflects the fact that mtDNA has a more rapid inter population divergence rate as a result of the lower effective population size for the mitochodrial genome. The way to understand this is that in a mating pair four different

copies of each autosomal nucleotide can be transmitted to the offspring but just one copy of mtDNA can be transmitted.

Interpretations of this greater intra population genetic diversity can actually support both the multi-regional and replacement hypothesis. If gene flow between two populations (continents) were large enough, then genetic homogeneity would have been maintained. This view is consistent with the multi-regional hypothesis. Genetic homogeneity in humans can also be explained if separation between the major continental populations (races) occurred recently with little time for divergence. This would support the African replacement hypothesis.

When DNA sequences are compared, the number of nucleotide differences between them can be counted. If two sequences differ by three nucleotides for example, then the inference can be made that at least three mutations have taken place since the two sequences were derived from a common ancestor. Making the assumption that natural selection is not a factor that effects that DNA sequence, then an estimation of a mutation rate allows for a date of this common ancestor to be inferred. It should then be possible to exam a large number of sequences for nucleotide differences and then work backward to find an ancestor from whom all existing variation started. This is called coalescence.

Because of its exclusive maternal inheritance, lack of recombination and relatively high mutation rate, mtDNA lends itself well to estimating the age of the most recent common female human ancestor. To assign a date, the rate of mutation per generation must be estimated. The mtDNA molecular clock has been calibrated by measuring the genetic differences between populations or species that have been

separated by a known length of time (chimps and humans diverged between 5 million to 9 million years ago). Application of the coalescent approach has led to of about 100,000 to 200,000 as the age of the most recent common female ancestor of living modern humans, or to be more precise, the age of the mitochondrial molecule ancestral to presently existing mitochondrial molecules is about 100,000 to 200,000 years (Cann, *et al*., 1987; Vigilant *et al*., 1991).

Genetic analysis of the Y-chromosome has placed the coalescence estimate of the ancestral male at about 188,000 years ago (Hammer, 1995). These result s lend credence to the selective neutrality for both mtDNA and the Y-chromosome data. It is unlikely that natural selection would effect two independent systems in the same manner.

While intriguing, these results must be taken lightly. Because mtDNA and the Y-chromosome are effectively a single non-recombining unit they are more subject to stochastic error. By taking into account the effects that stochastic error might have on the inheritance of mitochondria or the Y-chromosome the date of coalescence can theoretically be pushed back to 700,000 years ago (Nei, 1989; Nei, 1992).

Despite this caveat the coalescence date can provide important insight into the origin of modern humans. For one thing, it tells us the effective size of the human population. The coalescent time of mtDNA in generations (as expressed as t) , is directly related to the effective female population size (as expressed as Nf) as $t=2Nf$. Using the estimated coalescence date for mtDNA of 100,000 to 200,000 years, the effective population size has been relatively small throughout history: only a few thousand breeding females (Jorde, 1998; Takahata, 1993). A larger population would have meant more maternal lineages and the elimination of all of them would have taken longer then

100,000 to 200,000 years. Such a small population would not not have been able to cover the old world and exchange enough migrants to maintain genetic uniformity.

This is inconsistent with the multi-regional hypothesis (Jorde, 1998; Rogers, *et al*., 1995; Takahata, *et al*., 1995; Harpending, *et al*., 1993).

Mitochondrial DNA analysis has been used to determine which of the major human continental groups (African, Asian, European) exhibit the most genetic divergence (the most ancient).

One study attempted to "root" a dendrogram made from human mtDNA sequence data (Vigilant, *et al.,* 1991). A dendrogram is a diagram that summarizes the extent of genetic diversity between populations. The confidence in the divergence pattern of a dendrogram can be increased by the addition of a root. A root, if placed reliably, can orient a dendrogram to show which population split took place first. While statistical methods have been used to place a root (**Fig. 4**; Jorde, 1998) based only on the population understudy it is preferable to define a root by use of an outgroup. An outgroup is a genetically distant group from the population under study. The position the outgroup joins the dendrogram is the root. For example, chimpanzees are often used as an outgroup in human population studies.

The above mentioned study (Vigilant, *et al*., 1991) placed the root within African population. This study has been criticized for not being complete and performing incorrect rooting methods (Maddison*, et al*., 1992; Templeton, *et al*., 1993). A subsequent reanalysis of these data using more appropriate methods have once again placed the root within Africa (Penny, 1995). These data are consistent with the African replacement hypothesis as opposed to the multi-regional hypothesis.

A study analyzing genetic diversity among various populations was conducted utilizing the HV1 sequence of the human mitochondrial control region. This study once again provided evidence that supported the African replacement hypothesis. It demonstrated that a higher level of heterozygosity (a measure of genetic diversity) existed for African populations then for Asian or European populations (**Table 3** and **Fig. 5;** Jorde 1998).

Table 3. Hetrozygosity of Human Mitochondrial Control Region Based on HV1 Sequence Analysis

	Africa	Asia	Europe
Heterozygosity	0.023	0.015	0.010

The data obtained from mtDNA analysis and other molecular genetic markers is very compelling evidence towards the African replacement hypothesis. However, other explanations for this data also exist. For example, it is possible that humans could have originated elsewhere (Asia) but a population bottleneck reduced genetic diversity. Another possibility is that Africa had a greater effective population size then did the others. What has been demonstrated is how mtDNA analysis is employed as a tool in the search for the origins of modern man.

Another area mtDNA analysis has and is being used in is research into the genetic diversity, origins and history of present day mankind. Studies conducted over the past decade have documented mitochondrial genetic diversity among current human groups

(Direnzo, *et al*., 1991; Merriwether, *et al*., 1991; Stoneking, *et al*., 1990; Stoneking, *et al*., 1991) . Research like this, involving mitochondrial population genetic analysis in humans has been applied to the history of modern human populations. Some examples are the analysis of genetic diversity of pygmy tribes in South Africa (Vigilant, *et al*., 1989), sequence variation in the mitochondrial control region of Aboriginal Australians (van Host Pellekaan, *et al*., 1998), Asians (Melton, *et al*., 1996), Europeans (Melton, *et al*., 1997a) and sub-Saharan Africans (Melton, *et al*., 1997b).

An excellent example as to how mtDNA analysis has been used to study of modern populations is to summarize some of the research that has been performed on the origin and colonization pattern of Native Americans into the New World. Some of the questions raised on this issue are: the number of migrations that took place into the Americas, whether a bottleneck formed during colonization and just when did colonization begin.

A 1992 paper based on mtDNA RFLP studies (Wallace & Torroni, 1992) presented data that favored the bottleneck theory and claimed that four or five founding genetic variants were present in the initial colonization. Mitochondrial DNA sequence data also was consistent with the presence of four mtDNA lineages but found too much diversity in Native Americans to support the bottleneck theory (Ward, *et al*., 1991; Horai, *et al*., 1993).

Other authors have conducted research that examined the number of founding New World lineages as a function of their presence in both Asia and the Americas (Torroni, *et al*., 1993). One research team found evidence of a fifth group of related mtDNA lineages (Bailliet, *et al*., 1994). This work was criticized on the basis that the

nucleotide change that characterized the fifth group could have been the result of a mutation or admixture (Torroni & Wallace, 1995). A more recent study conducted on 108 individuals found in a 700 year old burial site concluded that the four major lineages where prersent bit also found evidence of at least one if not two more lineages (Stone, *et al.*, 1998).

Along with the controversy over the number of Native American mitochondrial lineages is the number of migrations into the Americas. This number has been reported as one (Bonatto, *et al*., 1997; Merriwether, *et al*., 1995), three (which was not based on mtDNA studies but linguistic analysis) (Greenberg, *et al*., 1986) or four (Horai, *et al*., 1993). The 1998 research team concluded that their data indicates just one wave of migration into the New World from Asia (Stone, *et al*., 1998). They based this conclusion on the fact that the lineages found in the Americas were relatively rare in Asia and is unlikely they would have been the only ones to enter in three or more migrations as opposed to coming over in one migration event.

The timing of the initial colonization into the Americas is yet another debatable issue. Archaeologists have placed it at 12,000- 20,000 years ago. Others believe it occurred greater then 20,000 years ago (Szanthmary, *et al*., 1993). Genetic data has provided evidence for colonization happening as late as 8,000 to 15,000 years ago (Ward, *et al*., 1991) or as late as 20,000 to 50,000 years ago (Forster, *et al*., 1996; Bonatto, *et al*., 1997; Torroni, *et al*., 1994). The study by Stone (Stone, *et al*., 1998) places the timing around 23,000 to 37,000 years ago.

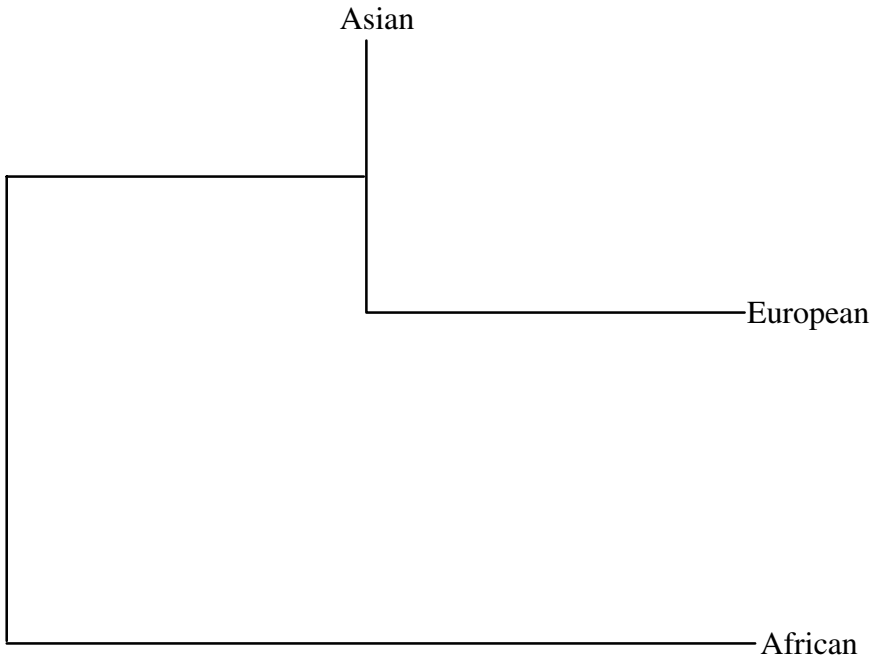


Fig. **4** A neighbor-joining dedrogram fro Africans, Asians, and Europeans based on (A) 411 bp of mtDNA fromhypervariable sequence 1 (HVS1).

Figure 5 illustrates an unrooted dendrogram of 15 populations living in Africa, Asia, and Europe. The results of this study again support the African Replacement Hypothesis with the African populatins grouping together and showing greater genetic distances from the Asian or African populations (Jorde, 1998).

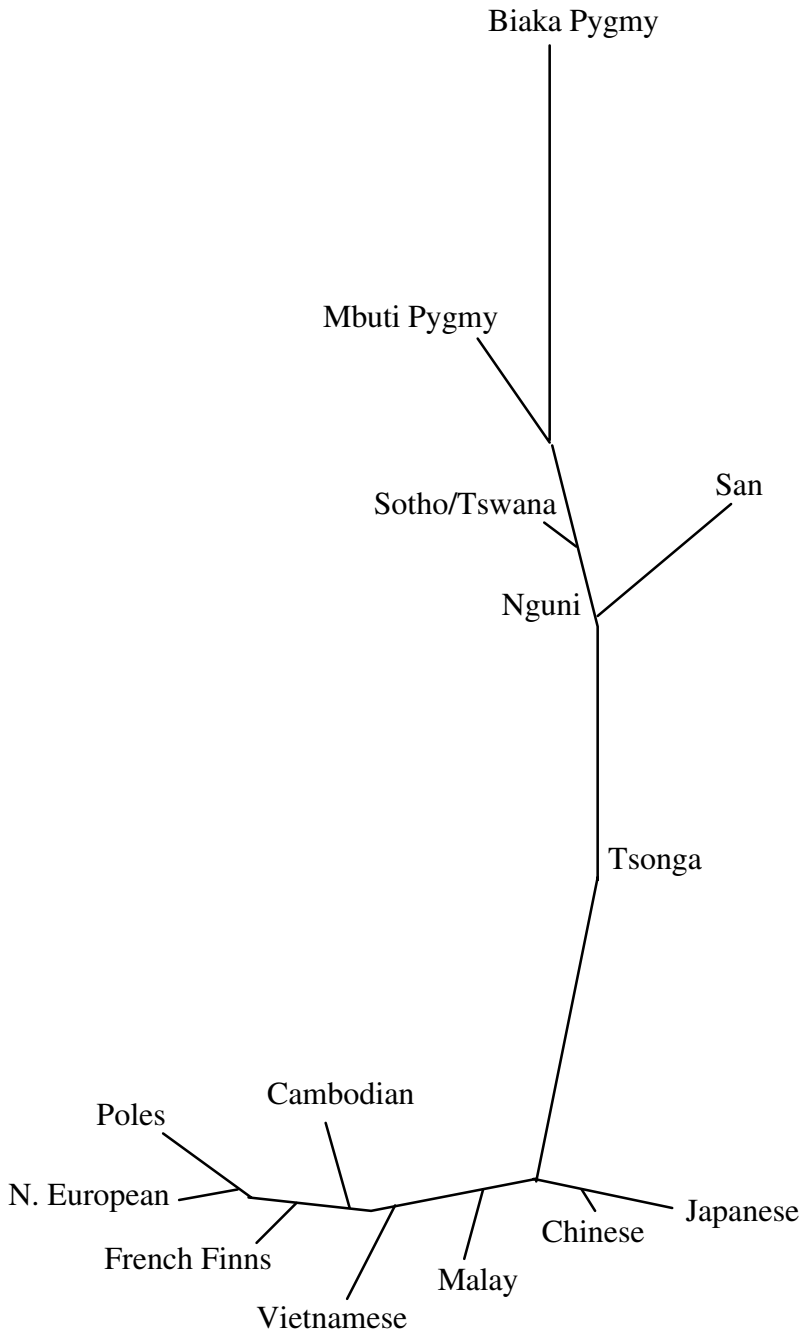


Fig. 5 An unroted network based on mtDNA (HSV1) from 15 populations living in Africa, Asia, And Europe. Genetic distances were estimated by using the Kimura 2-parameter model, and the neighbor-joining method was used to make the network.

Ancient DNA Studies Involving Human Mitochondrial DNA

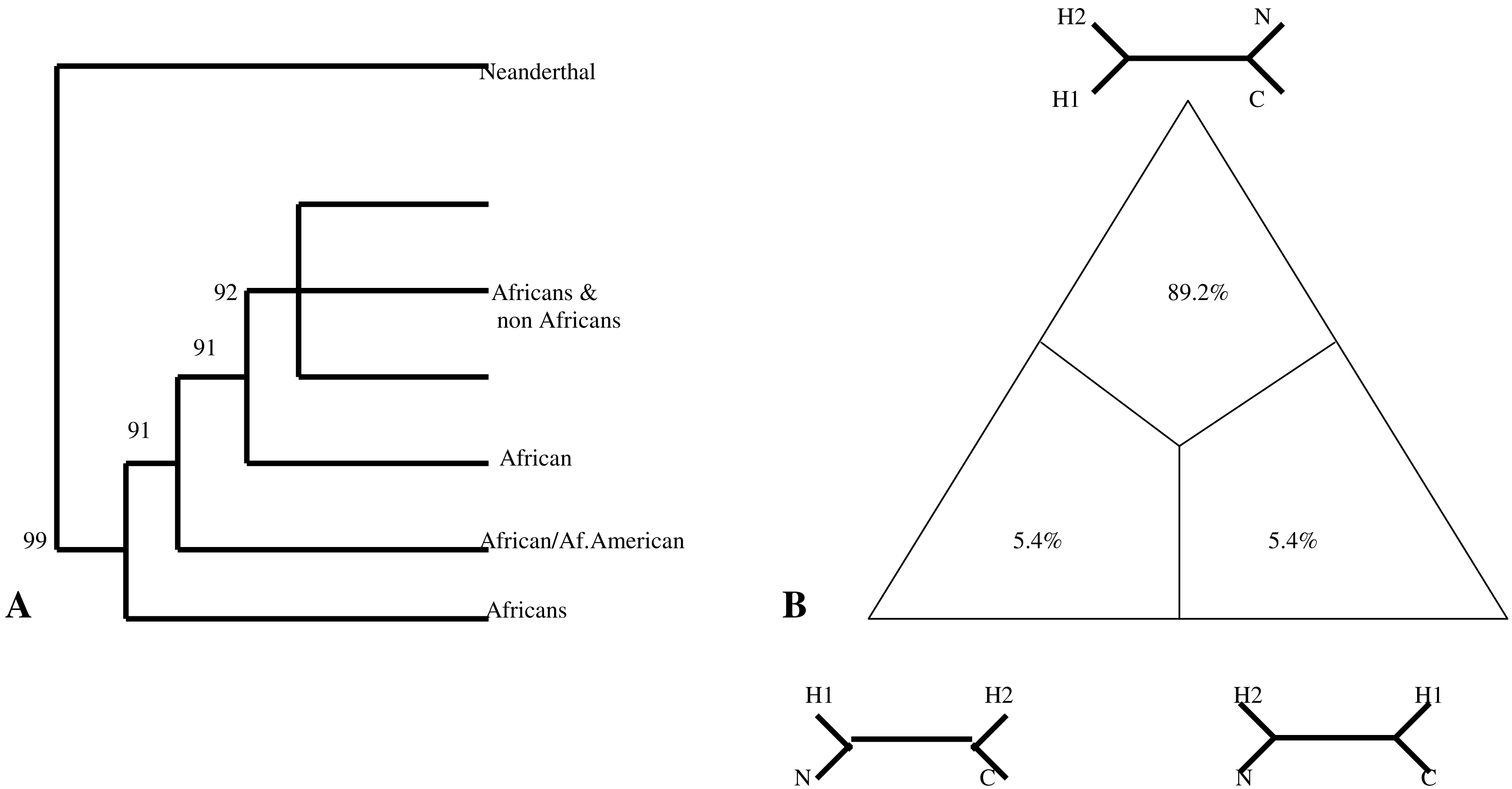
Polymerase Chain Reaction (PCR) technology has made it possible to retrieve interpretable DNA sequences from fossil specimens and to allow for comparison of that data with contemporary DNA sequences (Paabo, *et al*., 1989b). This methodology has been used on both human and non-human specimens. (Higuchi, *et al*., 1988; Paabo, 1989a; Rogan & Salvo, 1990a; Rogan & Salvo, 1990b; **Table 4**). Because of its high copy number per cell, small size, and amount of sequence data, the mtDNA genome is the primary molecule used in these types of studies. An example of this type of analysis was the recovery, amplification, and sequencing of mtDNA region V from a 7,000 year old human brain recovered in South Florida (Paabo, *et al*., 1988). Two mutations found in this sequence of mtDNA revealed that this particular haplotype was not seen before in American Indians, and thus described a third distinct maternal line. A more recent study (Krings, *et al.,* 1998) was performed on a Neanderthal skeletal specimen where a series of overlapping, short (100 base pair) PCR products from the control region were cloned and sequenced. A phylogenetic analysis was performed, and the Neanderthal sequences grouped as a separate branch unto themselves and did not group among modern humans. This lends credence to the theory that Neanderthals became extinct without contributing DNA to modern humans (**Fig. 6**). These are just a few examples of the various research studies involving mtDNA that have been and are continuing to be used to elucidate anthropological and evolutionary data on the human species.

Table 4. Cloned, Amplified, or Isolated Ancient or Extinct DNA Sequences.

Organism	Sequence	Age (Yr B.P.)	Reference
Human	Mitochondrial Region V	500, 800 300 – 750	Salva et al., . (1989) Hagelberg et al., , (1989)
		7,000	Paabo et al., . (1988)
	Alu repeated sequence	500, 800	Rogan and Salvo (1990)
		5,000	Paabo (1986)
	18S rDNA	500 – 2,000	Rogan and Salvo, in prep.
	28S rDNA	500 – 2,000	Rogan and Salvo, in prep.
	Mito. NADH dehydrogenase	300 – 750	Hagelberg et al., . (1989)
	Mito. D-loop	4,000	Paabo (1989a)
	Mito. cytochrome B2	7,000	Paabo (1988)
	Mito. Positions: 13282-13286	7,000	Paabo (1988)
Wooly Mammoth		40,000	Higuchi and Wilson (1984)
Quagga	Mito. cytochrome oxidase	140	Higuchi et al., . (1987)
	Mito. NADH dehydrogenase		
Marsupial wolf	Mito. 12S rDNA	120	Thomas et al., . (1989)
	Mito. cytochrome B		
Ground Sloth	Mito.	18,000	Paabo (1989a)
Z.mays	Mito. Cyto. C oxidase	1,500	Rollo et al., . (1988)
	H2a repeated sequence		
	Ribosomal RNA	1,000	Venanzi and Rollo (1990)
Cress	Ribosomal RNA	3,300	Venanzi and Rollo (1990)
E.virginensis		1,200	Rogers and Bendich (1985)
E. minutiflora		500	Rogers and Bendich (1985)
L.shockleyi		1,200	Rogers and Bendich (1985)
J.osteosperma		3,500->45,000	Rogers and Bendich (1985)
O.ramosissima		11,000	Rogers and Bendich (1985)
Symphoricatpos		>45,000	Rogers and Bendich (1985)
Magnolia	Chloroplast rbcL gene	>17-20 myr	Golenberg et al., . (1990)

Adapted from Rogan & Salvo

Fig. 6. A schematic phylogenetic tree relating the Neanderthal mtDNA sequence to 986 modern Human mtDNA sequences and likelihood mapping analysis showing the support for various groupings of Neanderthal, Human, and chimpanzee sequences.



(A) The tree was rooted with 16 chimpanzee mtDNA lineages. For clarity, only the first five branches without their internal branching structures but with their geographical states are shown. Numbers on internal branches refer to quartet puzzling probabilities. To calculate these, all possible combinations of the Neanderthal sequence, one of 16 chimpanzee lineages, and two of 100 lineages chosen at random from among 986 human lineages were analyzed. (b) The percentage of the quartets favoring the grouping of the Neanderthal sequence with the chimpanzee to the exclusion of the two human lineages is found in the upper of three areas. A total of 40 such analyses with different random sets of human mtDNA lineages were carried out and the average of these is given in the tree. The other internal branches were similarly analyzed.

The Use of Human Mitochondrial DNA in Human Identification and Forensic Testing

Recently, human mitochondrial DNA analysis has been used in the identification of human remains and forensic casework samples. Whether to assist in identifying unknown skeletal remains or in associating hair samples found at crime scenes, mtDNA is now a reliable court-accepted technique (personal communication, Stewart, 1999). Like the ancient DNA studies, the forensic use of mtDNA exploits its high copy number per cell, particularly in tissues ,such as hair, bone, and teeth),where genomic DNA analysis may not be possible (Higuchi, *et al*., 1988; Holland, *et al.,* 1995; Wilson, *et al*., 1995a). Another aspect of human mtDNA that makes it useful for forensic identification is its pattern of maternal inheritance,allowing, any known maternal relative (mother, grandmother, aunt, uncle, sibling, cousin, etc.) to be used as a reference sample.

The areas of the mtDNA sequence examined are Hypervariable Region I (HVI) and the Hypervariable Region II (HVII) found in the D-loop or Control region. They are short, approximately 300 to 350 base pairs each, and exhibit a high degree of variation between individuals. Sequences are analyzed by comparing differences in questioned samples to a known reference sample, the Anderson sequence (Anderson, *et al*., 1981). Guidelines for the use of mtDNA in forensic analysis have been suggested (Wilson, *et al*., 1993). Recently, the National Institute of Standards and Technology has developed human mtDNA standard reference material for increased quality control of human identity testing (Levin, *et al*., 1999). One well-known example of how mtDNA analysis can be applied to identification of unknown remains involves the last royal family of Russia, the Romanovs. In 1991, nine skeletons were excavated in Ekaternbuy, Russia.

Anthropological analysis tentatively identified them as Tsar Nicholas, Tsarina Alexandra, and their children. mtDNA analysis was used to confirm these findings by comparing mt DNA sequences from the remains to those of known living maternal relatives of the Tsar and Tsarina (Gill, *et al*., 1994; Ivanov, *et al*., 1996). The mtDNA analysis revealed that the sequences obtained from the adult female remains matched those of the remains of the children. The sequences of the adult maile matched those of the maternal descendent of the Tsar, with an exception of heteroplasmy at one base. When the mtDNA sequences of the adult male were compared to sequences obtained from the known remains of the Tsar's brother, they matched at all positions, including the shared heteroplasmy. Thus, the remains were confirmed as the missing Tsar and his family.

Dissertation Research

The information provided in this introduction illustrates how the basic biological properties of the human mitochondrial genome can be exploited in a number of interesting ways. The medical applications are shown to be immensely helpful in identifying what were at one time diseases of unknown origin. The use of mtDNA in anthropological and evolutionary studies answers many questions concerning human origin and the origin of populations and, perhaps more importantly, poses many questions and opens avenues for further research inquiry. In human identification, there are many instances where the remains are not amenable to analysis by the more conventional methods of anthropology and forensic dentistry. These cases include bone or teeth remnants that are either too small or too badly damaged for any analysis to be attempted other than mtDNA analysis. As previously mentioned, if a maternal living relative can be

found, that person's mitochondrial DNA sequence can serve as an excellent source of reference to compare against the sequence of the unknown evidentiary sample. To that end I have designed and performed two research projects to help improve the use of mtDNA analysis for forensic identification. To pursue this goal, I have studied the use of human mitochondrial DNA found in the D-loop region from two different perspectives. The first perspective was to design a series of experiments in order to determine what, if any, environmental insults would inhibit the application of human mtDNA from the D-loop region. Forensic samples are often subjected to a wide variety of exposures to many substances and conditions. This study was an attempt to simulate some of these conditions to determine what effect they would have on obtaining a PCR product. The second perspective was to determine the intergenerational mutation rate is for the D-loop region of human mtDNA. With the increased use of mtDNA in evolutionary studies and forensic analysis, it has become apparent to many investigators that a conflict exists between mutation rates determined from a variety of phylogenetic analysis studies and field observations from scientists involved in human identity testing. I performed a pedigree study using samples obtained by human volunteers who were members of families with multigenerational maternal lineages.

The reason for these two studies is that in cases involving human identification, much of the time, the samples that are received in the laboratory have been subjected to conditions that most research laboratories cannot duplicate. These conditions might include exposure to chemicals that might be used by a criminal to clean up a crime scene and remove evidence, such as water, soap, bleach, etc. Other conditions are found in nature, such as exposure to sunlight, soil, excess heat and/or humidity. The PCR

technique is a very robust one, yet the technique does have its limitations. The DNA polymerase, Taq polymerase, can be inhibited by certain substances such as hemin, a breakdown product of hemoglobin, or by the amount of DNA template used in the PCR procedure. The first study attempts to categorize what kinds of environmental exposures inhibited the PCR reaction and what treatments can be performed in order to optimize the PCR reaction. The second study proves useful in determining how often deviations might be seen in the mtDNA sequence between maternal relatives. For instance, if a sequence between a reference sample and an evidentiary sample is off by one base pair, should an analyst automatically say "No, this came from a different person?" or can the analyst take into account intergenerational mutations that might occur? The only way to answer this question is to obtain a better understanding of how often mutation rates occur within generations of the same family and eventually attempt to determine what mutations will occur or where they will occur within the D-loop regions.

Pedigree Analysis Studies Used in Determining Mutation Rate of Human Mitochondrial DNA in the Control Region

Sequence information obtained from the human mtDNA control region has been used increasingly by evolutionary biologists, molecular anthropologists, and systematic biologists. This information is used to elucidate such issues as the geographical origins of *Homo sapiens*, time of divergence from the most recent common ancestor of modern humans, and population bottlenecks and origins of Native Americans .(Di Renzo, *et al*., 1991; Horai, *et al*., 1993; Martin, *et al*., 1995; Sajantilla, *et al*., 1995; Stoneking, *et al*., 1992; Ward, *et al*., 1981). Polymerase Chain Reaction (PCR) technology combined with autoDNA sequencing has resulted in the increased use of mtDNA sequence analysis as an effective and reliable tool. Ancient DNA studies, identification of unknown human remains, and the forensic examination of hair samples are all areas which have benefited from the use of mtDNA sequence analysis (Higuchi, *et al*., 1988; Holland, *et al*., 1990b, 1995; Rogan & Salvo 1990a, 1990b; Wilson, *et al*., 1995b).

Because of the increased use of mtDNA sequence analysis, a large pool of data is now available for examination. It is becoming increasingly clear that some of the long-held assumptions about mtDNA may no longer be valid. One of these assumptions now in question is that mtDNA has a fixed mutation rate (Merriwether, *et al*., 1991; Stoneking, *et al*., 1992). The most commonly used approach to estimate the rate of human mtDNA evolution has been phylogenetic analysis also known as the relative branch length method (Cann, *et al*., 1987; Chen, *et al*., 1995; Horai, *et al*., 1995; Merriwether, *et al*., 1991; Ruvolo, *et al*., 1996). This method relies on obtaining mtDNA nucleotide sequence information from individuals and using it to construct a phylogenetic

tree via several analytical methods such as maximum parsimony, neighbor joining or maximum likelihood (Chen, *et al*., 1995; Hasegawa, *et al*., 1991; Horai, *et al*., 1992, 1995). An estimate is then made of the number of mutations that have occurred along each branch (this method requires a number of assumptions). In order to date the branch length of the tree, reference is made to a molecular clock that has been calibrated with divergence data from known archaeological or fossil records. Major assumptions of the phylogenetic analysis approach are a neutral evolution rate and a fixed rate of nucleotide substitution throughout the mitochondrial genome. In addition a relationship between sequence divergence and the accumulation rate of nucleotide substitution is assumed.

Early studies involving the use of mtDNA sequences estimated a single substitution rate while attempting to correct for multiple substitutions at the same site. It is now clear that this picture is not so simple (Cann, *et al*., 1987). For example, when the phylogenetic approach was used on the control region of human mtDNA, a sequence divergence rate of 7% to 22% per million years (myr) was found (Horai, *et al*., 1995; Pesole, *et al*., 1992; Stoneking, *et al*., 1992; Tamura & Nei, 1993). However, the same approach also found a divergence rate of approximately 2% to 4% per myr in coding regions of the mtDNA genome (Kocher, *et al*., 1991; Torroni, *et al*., 1994).

The presence of the HVI and HVII regions in the D-loop causes major problems in using the phylogenetic approach to determine divergence rates. These regions tend to throw off accurate estimates of branch length thus leading to inaccuracies in divergence rates (Penny, *et al*., 1995; Wakeley, 1993; Yang, 1995). Several other studies have taken these and other parameters into account in order to accommodate the various rates of substitution that occur throughout the human mitochondrial genome. These studies have

also attempted to take into account different classes of mutations (such as transitions versus transversions) (Adachi, *et al*., 1996; Hasegawa *et al*., 1991, 1993; Horai, *et al*., . 1992, 1995; Ruvolo, *et al*., 1993; Lundstrom, *et al*., 1992; Pesole, *et al*., 1992; Tamura, & Nei, 1993; Wakeley, 1993).

The results of recent studies have revealed that the rate at which mtDNA substitutions accumulate (over time) is a more complex process than previously imagined. For example, depending on which model for sequence evolution is used for the examined site on the mtDNA, the estimated range for the most recent common ancestor of humans is approximately 70,000 to 600,000 years ago (Adachi, *et al*., 1996; Cann, *et al*., 1987; Willis, 1995). Confidence intervals are not included in this range. Thus, the rate and patterns of mtDNA nucleotide substitution are still unclear.

It is commonly assumed that a child and his or her siblings should have exactly the same mtDNA sequence as their mother and all other maternal relatives. This assumption is the basis for much of the human identity and forensic testing being performed today. However, if a higher mutation rate for mtDNA exists, there may be situations where this assumption would prove to be incorrect. In fact, reports have indicated that this assumption is fallible. Several studies have shown that mutations can occur intergenerationally within various families of extended maternal lineages (Comas, 1995; Howell, *et al*., 1996; Parsons, *et al*., 1997; Wilson, 1997). Thus, the assumption that the mtDNA sequence remains stable in the maternal lineages of families must now have an added caveat that mutations, especially in the Control Region do occur. Along with the uncertainty of mutation rates and DNA divergence rates of the mtDNA genome as revealed in the previously cited phylogenetic studies (Holland, 1999; Parson, *et al*.,

1997), has led to an increased need to obtain a more clear-cut and accurate picture of mutation rate found in the human mtDNA control region.

An accurate assessment of the mutation rate in the mtDNA Control Region is essential, especially in cases of human identity or when forensic comparisons are being made. For instance, can a forensic examiner definitely exclude an individual from membership of a certain family based on one or two nucleotide differences in the sequences of the unknown and the reference sample? The following parameters were examined:

1. Presence of Mutations
2. If mutations were found, how were they categorized (single point mutations, insertions/deletions, heteroplasmy)?
3. If found, were the point mutations transition or transversion?
4. At what base pair position(s) were the mutations located?

Based upon the answers given in 1 through 4 above, the following questions must be asked: could the locations of these mutations be compared to others cited in the literature in order to determine if any locations were "hot spots" for mutations?, and could a determination be made of the mutation rates for this study and then compared to the mutation rates found in similar studies and those found using phylogenetic analysis? Due to prior research on mutation rate determination (Parsons, *et al*., 1997; Howell, 1996) a discrepancy in the interpretation guidelines between the FBI and AFIP laboratories has arisen (personal communication T. Parson and B. Budowle, 1999). The FBI laboratory uses a two base cut–off in making a decision as to whether to call the results a cannot exclude or exclude. The AFIP uses a three base cut-off level. Data

obtained by comparing sequence differences between the different lineages examined in this thesis will be used to see how well the two different laboratory criteria work and which method is preferable for casework in other laboratories. These data and their implications are presented in Results and detailed in the Discussion section.

CHAPTER II

MATERIALS AND METHODS

The methodology used in this study was to obtain DNA samples from buccal cell swabs from various individuals within several extended families, each family containing multiple maternal generations. The number of maternal generations ranged from two to four per familly. There were two to fourteen individuals per family for a total of 35 people. Each person was considered a separate mitochondrial generational event. The DNA from each subject was extracted and the HVI and HVII regions were amplified. The amplified DNA was sequenced, and the sequences were edited, aligned and compared, both to the Anderson reference sequences and to each individual within his or her respective family. A total of approximately 610 bases were analyzed between the HVI and HVII regions. Samples were sequenced for the forward and reverse strands. Any possible mutations were re-amplified and sequenced.

Human Subjects

Buccal cell swabs were voluntarily obtained from the 35 individuals representing five different maternal lineages. Each individual was instructed to swab the inside of the cheek for a period of 10 to 30 seconds with a sterile (*uv* treated) cotton applicator stick. The swabs were air-dried overnight and packaged in paper envelopes. The swabs were stored at –20°C until DNA extraction.

DNA Extraction

The cotton tip was removed from the applicator with sterile forceps and scalpel, and placed in a sterile 1.5 ml centrifuge tube. One ml of distilled water was added. The sample was incubated at room temperature for 30 minutes and then centrifuged at 10,000 to 15,000 x g for five minutes. The cotton tip was left in the tube and all but 30 µl of the supernatant was removed and discarded. Total cellular DNA was extracted using Chelex™ 100 resin (Walsh, *et al*., 1991). Chelex™ is a chelating resin that has a high affinity for polyvalent metal ions. It is composed of styrene divinylbenzene co-polymers containing paired aminodiacetate ions that act as a chelating group. A 5 % solution of Chelex™ was added to a final volume of 200 µl and incubated at 50°C for 30 minutes. The sample was then vortexed for 5 to 10 seconds and placed in a boiling water bath for eight minutes. It has been postulated that the boiling in the presence of Chelex™ prevents DNA degradation by chelating metal ions that act as a catalyst in the breakdown of DNA at high temperatures in low ionic strength solutions.

After the sample was removed from the boiling water bath it was once again vortexed for 5 to 10 seconds and centrifuged for three minutes at 10,000 to 15,000 x g. The resultant sample consists of denatured DNA.

DNA Quantitation

DNA quantitation was performed with a Slot-Blotting Manifold (Life Technologies, Inc., 1995) and chemiluminescence detection (ACES 2.0 Human DNA Quantitation System, Life Technologies, Inc., 1995).

The DNA was denatured by the addition of 100 µl of 0.5 M NaOH and 0.5 µl of NaCl to 2 µl of the sample. This was incubated for five minutes at room temperature. Simultaneously, a strip of Biodyne® A nylon membrane (Life Technologies, Inc., 1995) was soaked in 2XSSC for five minutes. The membrane was then placed in the Slot-Blot Manifold system, and the samples were pipetted onto it. A vacuum was applied for five minutes. The membrane was removed and rinsed in 0.2 M Tris-HCl and 2XSSC for a period of five minutes. The membrane was exposed to *uv* radiation for 90 seconds. A human specific DNA probe (D17V1) was hybridized onto the membrane at 50°C. After hybridization and washes, Lumi-phos® Plus (Life Technologies, Inc., 1995) was applied to the membrane. The membrane was encased in a static-free plastic folder and placed in a film development cassette. It was incubated for approximately 18 hours. Kodak X-0MAT AR film was exposed to the membrane for 15 minutes and developed. The intensity of the reaction was compared to the intensity of the known DNA quantitation standards.

DNA Amplification

DNA was amplified using the PCR (Polymerase Chain Reaction) methodology (Saiki, *et al*., 1988), followed by two different cycle sequencing methodologies. Two different amplification primer sets were used (**Table 5**).

The initial method used was the direct PCR methodology followed by the dye primer cycle sequencing with universal primers (Perkin-Elmer, 1995). This is a straightforward protocol involving the amplification of the entire Hypervariable regions in two separate reactions with "tailed" primers.. The primer tails are located on the 5' end and consists of universal sequencing primer sites (–21M13 or M13 Reverse). This allows both DNA strands to be sequenced separately using the same PCR product. Each amplification generates a 400 base pair product (**Table 5**).

Table 5. Amplification Primers for Dye Primer Sequencing

PCR	Primer Position	Sequence	Region Amplified
1F	M13FH16401	5'- TGT AAA ACG ACG GCC AGT TGA TTT CAC GGA TGG TG-3'	HVI 15975-16420
1R	M13RL15996	5'- CAG GAA ACA GCT ATG ACC CTC CAC CAT TAG CAC CCA AAG-3'	
2F	M13RL00029	5'- CAG GAA ACA GCT ATG ACC GGT CTA TCA CCC TAT TAA CCA C-3'	HVII 00008-00429
2R	M13FH00408	5,- TGT AAA ACG ACG GCC AGT CTG TTA AAA GTG CAT ACC GCC-3'	

note- base pair positions refer to the Anderson Sequence
note- M13 universal primer sequences are underlined

The PCR Reaction Mix is made up of 1.25 mM of dNTP, 10 XPCR Buffer (100 mM Tris HCl pH9.0 at 20°C, 15 mM MgCl2, 500 mM KCl, 1 mg gelatin per ml and 10 mg BSA per µl (molecular biology grade), 5pMol PCR primers per µl , 5 U/µl AmpliTaq™ DNAPolymerase (Perkin-Elme, 1995) and 5 ng of DNA (**Table 6**).

Table 6. Reaction Mixture Protocol (Dye Primer PCR Amplification)

Reagent	**Volume**	**Concentration**
DNA	q.s.	> 50 ng
Forward primer	2 µl	10 pmol
Reverse primer	2 µl	10 pmol
dNTP mix	4 µl	50 uM
10X PCR buffer	10 µl	1X
BSA	1 µl	10 mg/mL
dH2O	q.s.	
Final Volume	100 µl	

The samples were amplified with a Perkin-Elmer 2400 thermocycler under the following conditions: The sample tubes were placed in the thermocycler which had been preheated 85°C for two minutes. AmpliTaq™ DNA Polymerase was then added. The amplification parameters were as follows: 94°C for 45 seconds, 66°C for 60 seconds, and 72°C for 60 seconds. This was repeated for 28 cycles, linked to a 4°C soak (**Table 7**).

The presence PCR product was verified using a 2% agarose 1 x TBE ETBR-stained gel. The PCR product was purified using Micrcon 100 centrifugation filtration to prepare the sample for cycle sequencing.

Table 7. Thermal Cycler Conditions (Perkin-Elmer 2400) Dye Primer Amplification

Temperature	Time	Number of Cycles
94^0 C	45 seconds	1 cycle
66^0 C	60 seconds	
72^0 C	60 seconds	28 cycle
4^0 C	Soak	∞

The second methodology employed overlapping PCR followed by DYE terminator cycle sequencing (Perkin-Elmer, 1995). This method is similar to the previous one in that each hypervariable region is amplified directly from the extracted DNA, however, each region is amplified in two smaller overlapping fragments (**Table 8**). The PCR products that are produced are each approximately 270 base pairs long with approximately 100 bases of shared sequence between them.

Table 8. Sequencing Primers for Dye Terminator Overlapping PCR of HVI and HVII

PCR/Sequence	Primer Position	Sequence
1 F	F 15971	5'-TTA ACT CCA CCA TTA GCA CC-3'
1R	R 16255	5'- CTT TGG AGT TGC AGT TGA TG- 3'
2F	F 16144	5'- TGA CCA CCT GTA GTA CAT AA-3'
2R	R 16414	5'- CAC GGA GGA TGG TGG TCA AG- 3'
3F	F 29	5'- CTC ACG GGA GCT CTC CAT GC- 3'
3R	R 270	5'- TGG AAA GTG GCT GTG CAG AC-3'
4F	F 155	5'- TAT TTA TCG CAC CTA CGT TC-3'
4R	R 381	5'- GCT GGT GTT AGG GTT CTT TG-3'

Primer set one of the HVI region spans 15,997 to 16,236 base pairs; primer set two spans16,159 to 16,391 base pairs. In HVII, primer set three spans 048-285 base pairs, and primer set four spans 172 to 408 base pairs.

The amplification master mix is the same as in the universal primer method, except it is carried out in a 25 µl reaction. Thermocycling was carried out using either on a Perkin-Elmer model 9600 or 9700 thermocycler. The parameters were 95°C for 60 seconds, followed by 32 cycles of the following: 95°C for 10 seconds, 60°C for 30 seconds, and 72°C for 30 seconds. This was linked to a 10 minute hold at 15°C followed by 4°C soak (**Table 9**). The resulting PCR product was purified by Microcon™ 100 centrifugation filtration.

Table 9. PE 9700 Thermal Cycler Profile Dye Terminator Cycle Sequencing Reaction.

Temperature	**Time**	**Number of Cycles**
96^0 C	15 seconds	25 cycles
50^0 C	5 seconds	
60^0 C	2 minutes	
4^0 C	Hold	

DNA Sequencing and Analysis

Two major cycle sequencing methods were used. The first one is dye primer labeling, In this method, the primers are tagged with four different fluorescent dyes in four separate base specific reactions.The second strategy was the dye terminator labeling method. This involved tagging each of the four dideoxy terminators with a different fluorescent dye. The growing chain is both terminated and labeled with a dye that corresponds with a specific base.

For the dye primer method, the ABI Prism Dye Primer Cycle Sequencing Ready Reaction Kit™ (PE Biosystems, 1995), was used. The kit contains four pre-mixed components. An **A** mix (ddATP, dATP, dCTP, 7-deaza-dGTP, dTTP, JOE dye primers, Tris-HCl (pH9.o0), MgCl2, thermal stable pyrophosphartase, and AmpliTaq™ DNA Polymerase), the **C** mix (ddCTP, dATP, dCTP, 7-deaza-dGTP, dTTP, FAM dye primer, Tris-HCl (pH9.0) MgCL2, Thermal stable pyrophosphatase, and AmpliTaq™ DNA Polymerase), **G** mix (ddGTP, dATP, dCTP, 7-deaza-dGTP, dTTP, TAMARA™ dye primer, Tris-HCl (pH 9.0), MgCl2, thermal stable pyrophosphatase, and AmpliTaq™ DNA polymerase), and **T** mix (ddTTP, dATP, dCTP, 7deaza-dGTP, dTTP, ROX™ dye primer, Tris-HCl (pH 9.0), MgCl2, thermal stable pyrophosphatase, and AmpliTaq™ DNA polymerase).

Cycle sequencing technologies employ the standard Sanger dideoxy chain termination method (Sanger, *et al*., 1977). By coupling this chemistry with a thermostable DNA polymerase, the sequencing reactions can be repeated on the same template in an automated cycling fashion. In each cycle, the primer is annealed to the template, the normal dideoxy reactions occur, the newly synthesized double-stranded DNA is denatured, and the cycling reaction is repeated (Alpney, 1997).

Cycle sequencing was performed on a PE 2400 (**Table 10**). The resulting products were purified by ethanol precipitation and resuspended according to manufacturer's protocol (PE Biosystems, 1995).

Table 10. Dye Primer Cycle Sequence Reaction Set-Up and Thermal Cycler 2400 Profile

Reagent	**Reaction**	**A**	**C**	**G**	**T**
PCR Template		1µl	1 µl	2 µl	2 µl
-21 M13 or M13 Rev		4 µl	4 µl	8 µl	8 µl
Final Volume		5 µl	5 µl	10 µl	10 µl

PE 2400 Thermal Cycler Profile

Temperature	**Time**	**Number of Cycles**
95^0 C	30 seconds	
55^0 C	30 seconds	15 cycles
72^0 C	1 minute	
LINKED TO FOLLOWING		
90^0 C	30 seconds	
72^0 C	1 minute	15 cycles

The dye terminator protocol was performed with an ABI Prism d Rhodamine Dye Terminator Cycle Sequencing Ready Reaction Kit™ (PE Biosystems, Perkin-Elmer Corp.). This kit contains terminator ready reaction mix (**A**-Dye terminator labeled with dichloro [RGG], **C**-Dye Terminator labeled with dichloro [ROX™], **G**-Dye Terminator labeled with dichloro [RHO], **T**-Dye Terminator labeled with dichloro [TAMRA}, deoxynucleoside triphosphates (dATP, dCTP, dITP, dUTP), AmpliTaq™ DNA polymerase FS with thermally stable pyrophosphatase, MgCl2 and Tris-HCl buffer pH 9.0). The sequencing reaction set-up is illustrated and carried out in a Perkin-Elmer 9700 Thermocycler used according to standardized protocol. The product was purified using ethanol precipitation according to manufacturer's specifications (PE Biosystems, 1995).

The resulting DNA sequences were read utilizing the ABI Prism™ 310 Genetic Analyzer (**Fig.7**). The ABI 310 Genetic Analyzer™ utilizes capillary electrophoresis to fractionate the fragments and detect the dye-labeled bases through laser excitation and spectral analysis (Perkin-Elmer, 1995; Butler, 1998; Zhang, *et al*., 1996). Tubes containing the samples are placed in a tray. Next, that tray is put on the 310's autosampler (**Fig.8**). The ABI Prism 310 Genetic Analyzer™ functions by first placing the sample tubes in a tray that are then placed in the instrument's autosampler. The autosampler brings each sample into contact with the cathode electrode and one end of a glass, polymer-filled capillary tube. The anode electrode at the other end of the capillary is immersed in a buffer.

A small amount of sample enters the capillary as the current flows from anode to cathode through electrokinetic injection. The end of the capillary, near the cathode, is then placed in a buffer. Current is once again applied and the electrophoresis continues.

When the fluorescent dye labeled nucleotides reach a detector window in the capillary, a laser excites the dyes. The emitted fluorescence is collected by a CCD camera (**Fig.9**). Software then interprets the results, calling the bases from a fluorescence intensity at each data point (Perkin-Elmer, 1995).

The resulting sequences were edited, aligned and compared using Sequentier 310 Software and GenWorks™ 2.51 Software (Perkin-Elmer, 1995).

Fig. 7. ABI 310™ capillary electrophoresis unit (front view).

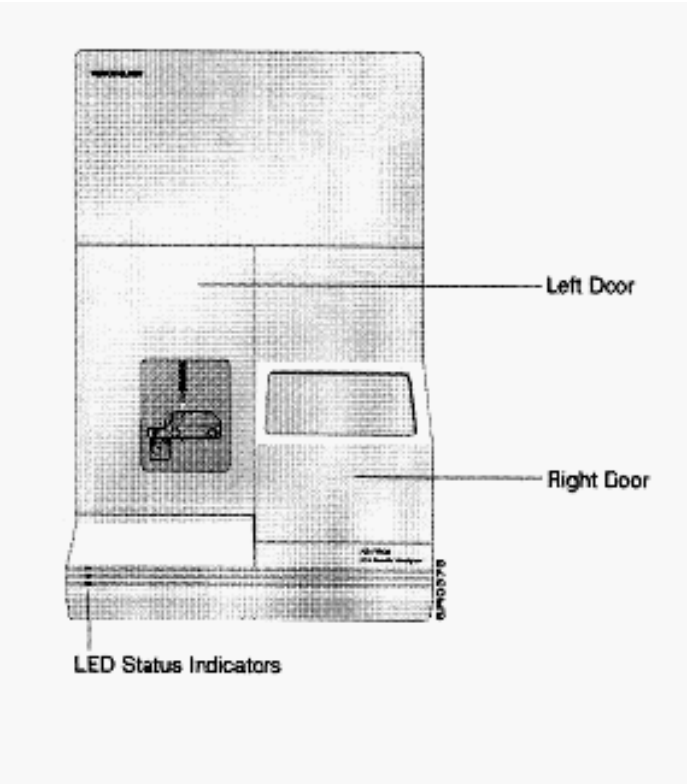


Fig. 8. ABI 310™ Capillary Electrophoresis Unit (Interior).

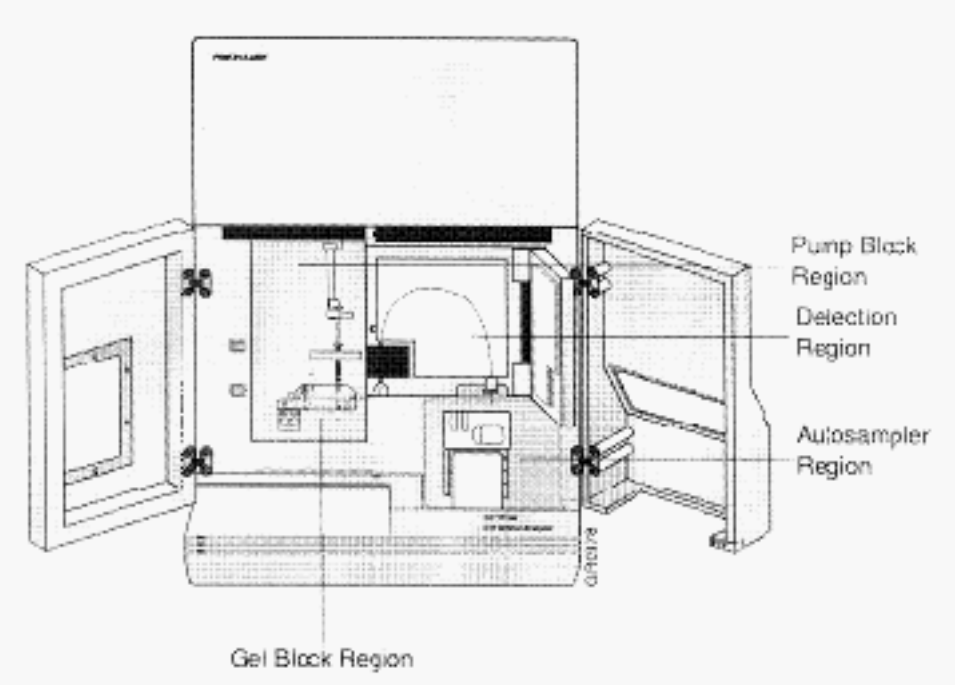
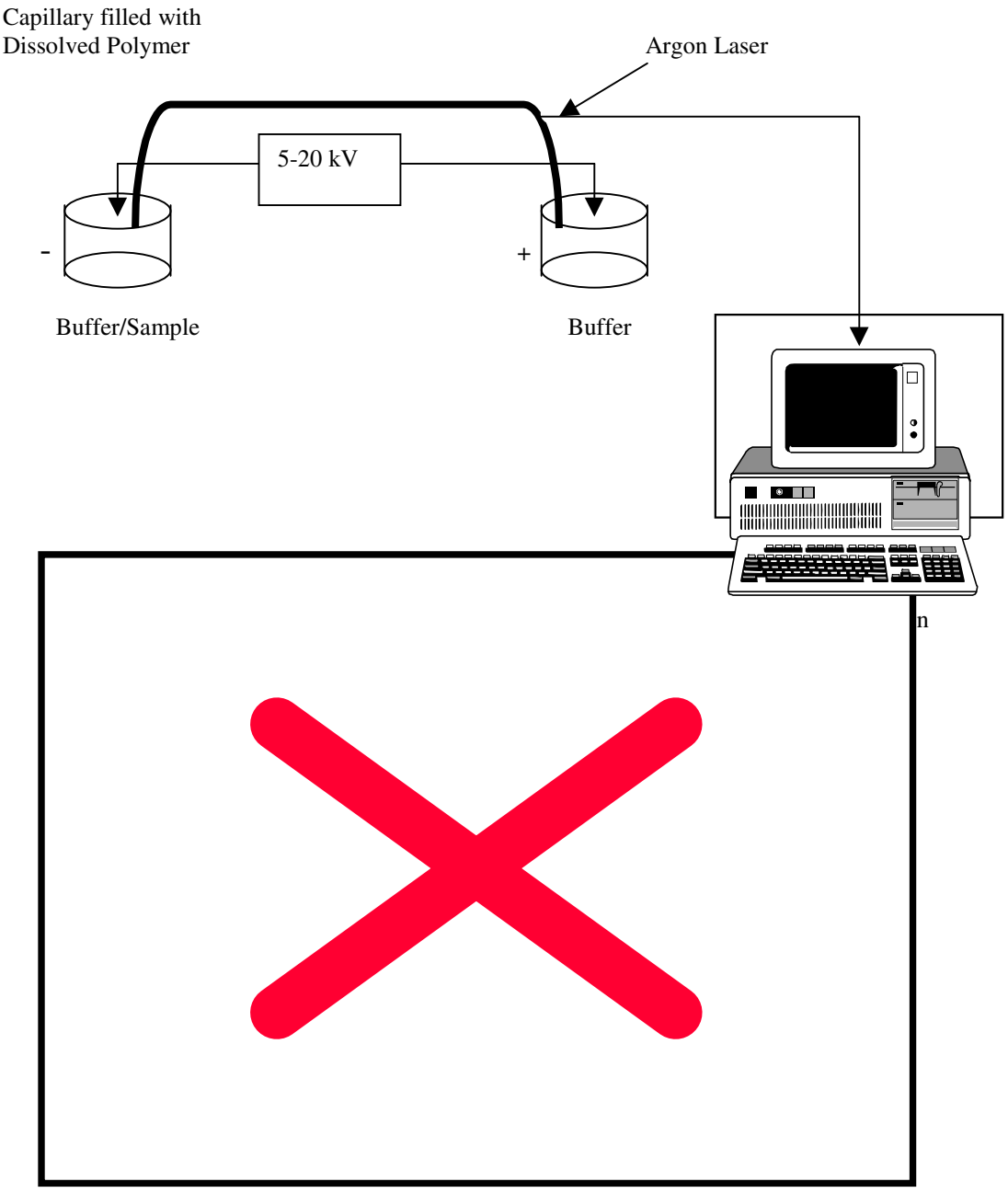


Fig. 9. Schematic of a CE Instrument with Laser-Induced Fluorescence Detection.



Results

Mitochondrial DNA from the HVI and HVII regions was sequenced and analyzed from five different families: PLI, SMI, RGI, RGII, and GBI. PLI had a total of 10 mitochondrial generations (each individual is considered a separate mitochondrial generational event) spanning four maternal `generations (Fig. 10, 11, and 12). The SMI family had a total of six mitochondrial generations. The RGI family had two mitochondrial generations and one maternal generation, and RGII had three mitochondrial generations within two maternal l generations. The GBI family was comprised of 14 mitochondrial generations found between three maternal generations. None of the families had any diagnosed pathologies associated to mutational events in coding genes of the mitochondrial genome.

SMI Family

The SMI family sequence analysis showed 100% concordance among all of the individuals tested. No mutational events were detected.

GBI Family

The GBI family had 100% concordance among all the individuals tested. No mutational events were observed. NOTE: An interesting sidebar to this particular family was that three of the individuals tested were triplets. It was not known which of the three were monozygotic twins and which one was not. A kinship analysis was performed looking at eight loci of chromosomal DNA. These eight loci were tetranucleotide repeats, commonly called short tandem repeats (STRs) (Perkin-Elmer, 1998; Edwards, *et al*., 1991, 1992; Wallin, 1998).

Figure 10. Pedigree Chart of GBI Family

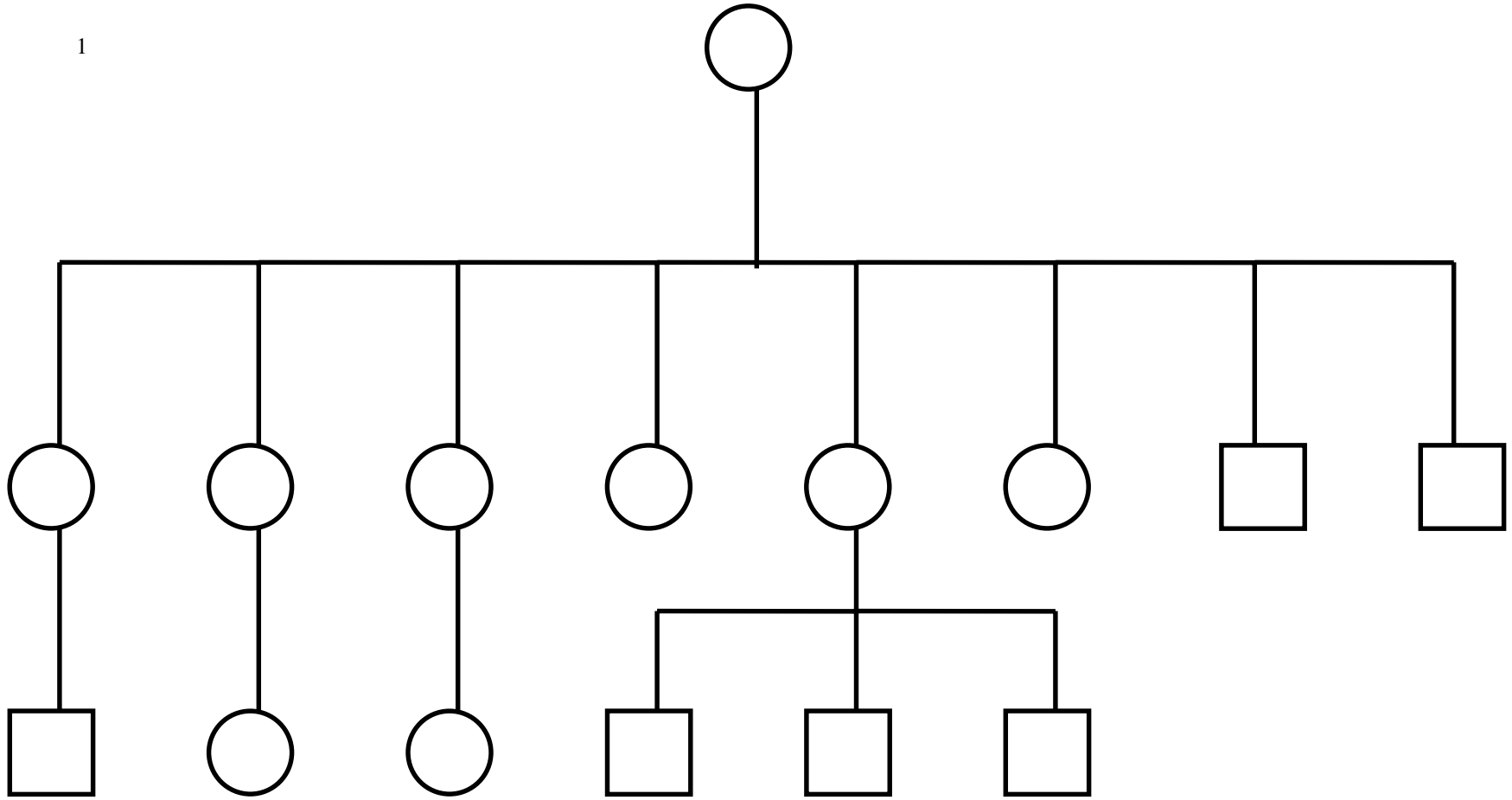
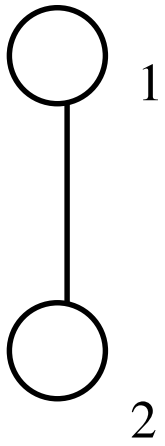
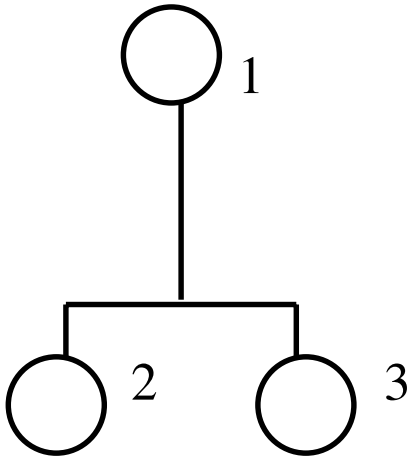


Figure 11. Pedigree Chart of RGI and RGII Families

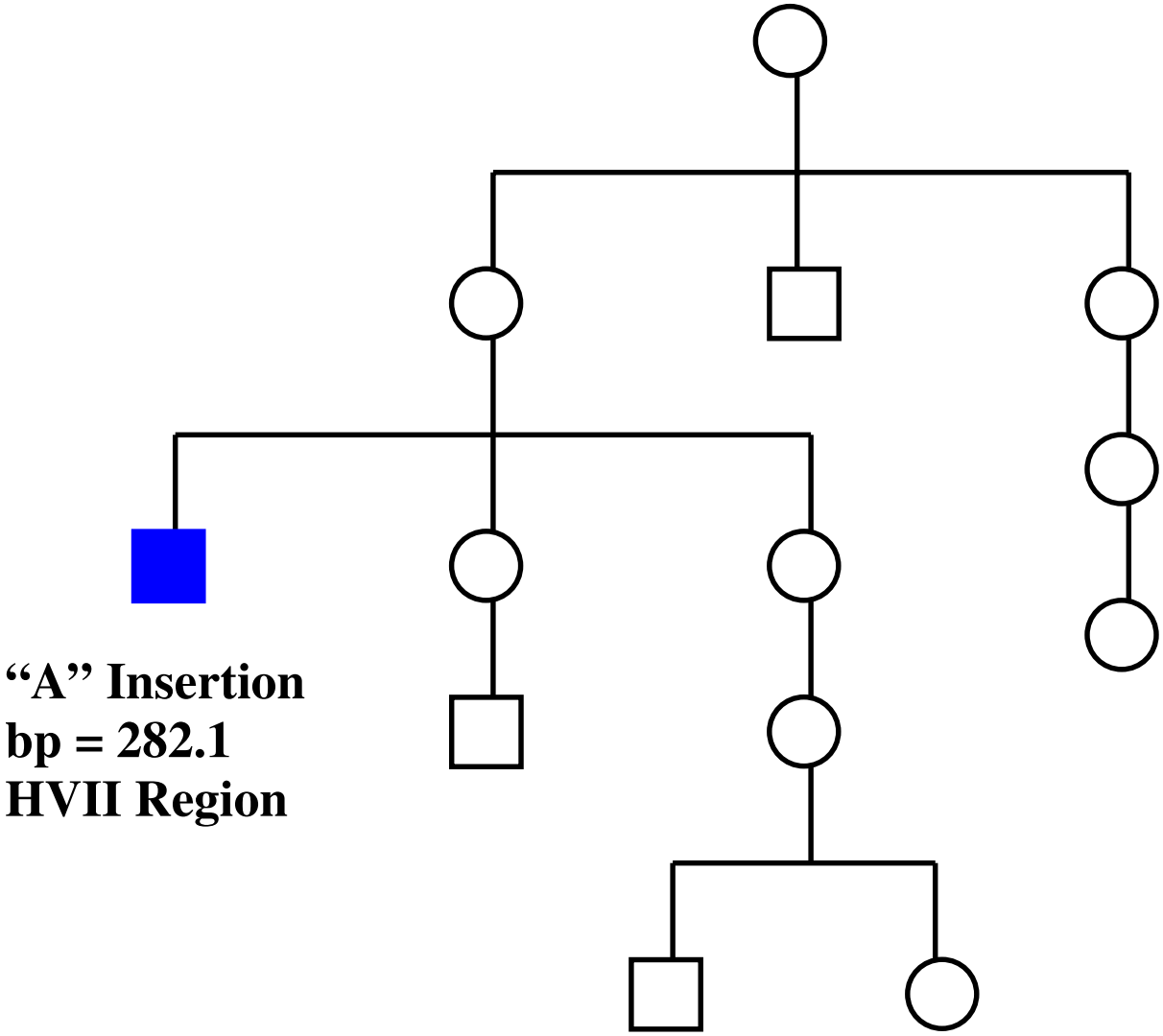


RGI



RGII

Figure 12. Pedigree Chart of PLI Family (Location and Type of Mutation)



The kinship analysis of the STR loci did prove and demonstrate which two of the triplets were monozygotic twins (**Table 11**).

Table 11. STR Results For Kinship Analysis Triplets In GB I Family

STR Loci	**Sister #1**	**Sister # 2**	**Sister #3**
D3S1358	15,16	15,16	15,16
VWA	17,18	17,17	17,17
FGA	21.2,22	20,21.2	20,21.2
Amelogenin	XX	XX	XX
D8S1179	12,13	13,13	13,13
D21S11	31,31.2	31,31.2	31,31.2
D18S51	16,18	15,18	15,18
D5S818	11,12	10,11	10,11
D13S317	10,11	10,11	10,11
D7S820	9,10	9,10	9,10

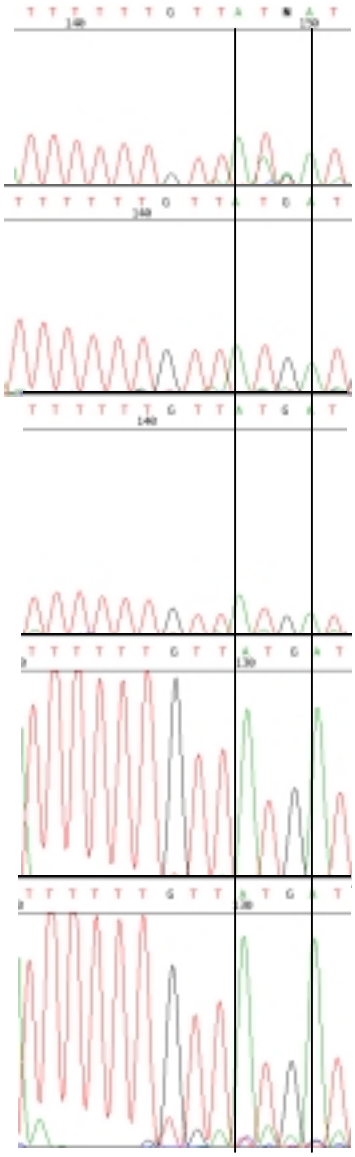
RGI and RGII Families

No mutational events were observed for the RGI and RGII families. All mtDNA sequences were in concordance with one another.

PLI Family

The PLI family failed to reveal any detectable mutational events on the HVI region for any of the individuals tested. One mutational event, an A insertion, was observed in individual number 4. This insertion event occurred at position 282.1 of the HVII sequence of the control region (**Fig. 13**).

Fig. 13. Electropherogram delineating mutational event at 282.1, HVII region.



Mutation Rate Analysis

The mutation rate was calculated using a simple method employed by Parsons, *et al*., (1997). There was one mutational event observed in 35 mitochondrial generations tested. Assuming a generation time of 25 years, this calculates out to one event every 875 years. Approximately 610 base pairs were analyzed per generation. Multiplying 875 years by 610 base pairs calculates to 533,750 possible mutations per site. Dividing this into one million years gives an estimate of 1.9 mutational events/site/myr.

Discussion

The results of this research coincide with those found in similar studies (Parsons, *et al*., 1997). After sequencing the D-loop regions of HVI and HVII in 327 individuals from the United States and Great Britain, Parsons found 10 instances of nucleotide substitution. This is approximately one mutational event in every 33 generations which extrapolates to a mutation rate of 2.5 substitutions/site/myr with a 95% Bootstrap Confidence Interval was 1.4 – 4.0/site/myr.

Another pedigree study by Howell, *et al*., (1996) analyzed the D-loop region in an extended family of 49 individuals and found two mutational events. This is a mutation rate of approximately one event in every 25 generations. Even if an assumption is made that the remaining maternal relatives whose D-loop regions were not analyzed would reveal no mutations (a conservative, reasonable assumption), then the mutation rate is still high, one in 40 mutations per generation. This study has been criticized because the family upon which the study is based had a history of LHON and thus might have been prone to mutational events in the mitochondrial genome.

Despite that particular criticism of Howell's paper, his results are in very close agreement with those of the Parson's study and those found in this dissertation research. A study by Bendal, *et al*., (1997), dealt specifically with the mutation rate of heteroplasmy. DNA from 180 twin pairs was sequenced at the D-loop region. Three of these pairs revealed heteroplasmy.

The combined weight of these studies reflects a significantly higher mutation rate than those derived by phylogenetic analysis. Phylogenetic analysis performed on the control region has derived a range of D-loop substitutions roughly between 0.025 to 0.26 sites/myr (Hasegawa, 1991; Horai, 1995; Tamura & Nei, 1993; Vigilant, 1991). The mutation rate reported in this dissertation (1.9/site/myr) is approximately 80 to 800 times higher than those reported in the above-referenced phylogenetic studies.

The reason for the disparity in the mutation rate analysis results between the two methods is still a matter of conjecture. However, it does lead to some interesting speculations. A list of compiled mtDNA mutations (**Tables 12 & 13**) reveals what appear to be several hot spots for mutational events in the HVI and HVII regions. Thus, the pedigree analysis approach might be detecting nucleotide substitutions at these hot spots while the phylogenetic analysis method rates are averaged over all the sites.

Table 12. Known Inter-Generational Substitutions and Heteroplasmy (HVI).

HVI

Position	Type
16092 C>T	Substitution
16093 C/T	Heteroplasmy
16093 C/T	Heteroplasmy
16093 T>C	Substitution
16093 T/C	Heteroplasmy
16129 A/G	Heteroplasmy
16169 C/T	Hetero/Substitution
16172 C/T	Heteroplasmy
16192 C/T	Heteroplasmy
16222 A/G	Heteroplasmy
16222 A>G	Substitution
16239 C/T	Heteroplasmy
16262 C>T	Substitution
16256 T>C	Substitution
16293 A>G	Substitution
16293 A/T	Heteroplasmy
16295 C/T	Heteroplasmy
16311 C/T	Heteroplasmy
16318 A/T	Heteroplasmy
16355 C/T	Heteroplasmy/Substitution

Table 13. Known Inter- Generational Substitutions and Heteroplasmy (HVII).

HVII

Position	Type
94 A>G	Substitution
152 T/C	Heteroplasmy
185 G>A	Substitution
189 A>G	Substitution
195 C/T	Heteroplasmy
199 C/T	Heteroplasmy
207 A>G	Substitution
207 G>A	Substitution
234 A/G	Heteroplasmy
282.1 insertion/deletion	Substitution
309.1 insertion/deletion	Substitution
309.2 insertion/deletion	Hetero/substitution

Several evolutionary studies indicate that, over long time periods, different sites exhibit highly different substitution rates (Hasagawa, 1992; Pesale, *et al*., 1992; Wakeley, 1993).

The pedigree approach naturally involves a very small period of evolutionary time. Therefore, this analytical method might be observing substitution rates where they occur most rapidly. However, if one examines **Tables 12**& **13** approximately 50% of the mutations occur at what appear to be hot spots. While mutational hot spots might offer a partial explanation, more studies are needed to explain the high mutation rate detected by pedigree analysis.

Another possible explanation of the differences in mutation rates derived from the evolutionary analysis and from pedigree analysis is that the evolutionary studies detects only the mutations that have become fixed in the population over time whereas the pedigree approach detects mutations that are occurring in an individual, but may not remain in the whole population. This makes sense in light of the fact that most of the evolutionary studies referred to use the phenetic approach to building phylogenetic trees. This approach depends upon measuring genetic distance as a function of sequence differences between species. This approach naturally involves long periods of time.

According to standard population genetic theory, the probability that a new allele will become fixed within a population is a function of the initial frequency of the allele, any selective value it possesses, and the effective population size (Li, 1998).

With mtDNA, the effective population size has two main determining factors. The first one is the number of mtDNA molecules in the gene line. The other is the population size and demographics of the population sample in question. Also, it is not

known what, if any, selective pressures exist for D-loop mutations. Some studies have suggested that mtDNA evolution is not strictly neutral (Howell, *et al*., 1997).

It is possible that the high mutation rate obtained from the pedigree studies, compared to the rate found in phylogenetic analysis, might be a reflection of the failure of many control region mutations to become fixed in the population. There are no definitive data on the fixation probability of control region mutations, but it is a safe assumption that it is less than 1. In other words , it is highly unlikely that 100% of all mutations become fixed. Taking this assumption into consideration, it is not surprising that the mutation rates derived from pedigree analysis are higher than those obtained by the phylogenetic method.

Because of the lack of information on the subject, the role played by selection or random genetic drift in the fixation process of mtDNA is open to question and skeptical inquiry. One observation made in the Parson study (Parson, *et al*., 1997) was that some of the control region mutations reported occurred at sites with below average levels of polymorphism within the population. Statistically, pedigree analysis should be a relatively unbiased indicator of mutation rates (Paabo, 1996) at the most rapidly mutating sites. If selection is not a major factor, these sites should correspond to the most rapidly evolving sites discovered using phylogenetic analysis. Based on the results of the Parson's research (Parsons, *et al*., 1997) this may not be so. Indeed, some negative selection against D-loop mutations might be taking place. Mutations that are fixed in an individual might not be fixed in a population over evolutionary time. Thus, the effect of random drift might be more of a factor in pedigree analysis than selection, which needs

longer time periods to have its effect felt, and it is the effect of selection that is detected by phylogenetic measures. However, more in-depth study is needed in this area.

The bottleneck theory states that a small subset of the mitochondrial genome in the female germ line is transmitted to her offspring. Consequently, a fraction of new mtDNA mutations will be transmitted to subsequent generations, but the fixation rate from the transmitted fraction will occur more rapidly if a bottleneck occurs (Howell, *et al*., 1996). Therefore, the mutation rate in mtDNA would be higher than the divergence rate, especially if a significant proportion of the mutations became fixed in an individual but not in populations.

Pedigree mutation rate analysis might provide some information of recombination, or even paternal contribution of mtDNA, as being responsible for some of these high mutation rates. As was mentioned in the introduction, there is some evidence that recombination is possible in mtDNA.

A recent study (Awadalla, *et.al*., 1999) found evidence of linkage equilibrium in hominid mtDNA. The reported research found evidence that linkage disequilibrium declines as a function of distance between sites. The researchers came to the conclusion that recombination, specifically recombination with paternal mtDNA is the most reasonable mechanism for this.

Another explanation for mitochondrial mutations might be an effect of time. A paper published in Science (Michikawa*, et al*., 1999) reports results that give a strong indication that mutations in the human mitochondria control region increases with age of the individual. The researches analyzed fibroblast cells from 18 people with ages ranging from 1 to 101 years. The research team also obtained two sets of stored cell pairs taken

9-19 years apart from 9 subjects. The control region was analyzed for mutations using Density Gradient Gel Electrophoresis and sequencing. Mutations were found to be present in5% to 50% of the clones from the older individuals and none in the younger ones. Analysis of the 9 cell pairs revealed that 3 of them had at least one mutation in the older cells but not the younger ones. It was interesting to note that one of the reported mutations was a "T" insertion approximately 100 bases downstream from where the mutation reported in this dissertation was found. This aging study was of particular interest given that the affected individual with the mutation found in this dissertation was the youngest of three siblings and the mother was of advanced maternal age (approximately 40 years).

No single cause can be directly pointed out as to being the primary mechanism for mitochondrial mutations, the fact that they do exist can be a starting point to a greater understanding of mitochondrial biology. While the mechanism for mtDNA mutations is as yet speculation what is not speculation is that the results of this and similar studies do indeed have to now be taken into account in evolutionary research and forensic analyses.

There are two main classes of mutations that can occur in mtDNA. One of them, sequence modifications, is the subject of this research. The second type of phenomena is called heteroplasmy. While heteroplasmy is not the main emphasis of this dissertation a discussion of it is essential in any paper dealing with mtDNA mutations and their effect on the interpretation of forensic human identity tests. Heteroplasmy can be defined as the existence of two or more subpopulations of mtDNA in one individual.

To understand how heteroplasmy or sequence modifications form in mtDNA requires some understanding of the biology of the system. During embryonic

development, mtDNA molecules are replicated independently of one another. They are not strictly tied to mitotic or meiotic cell division. Mitochondrial DNA replication is also associated with a much higher error rate than nuclear DNA (Brown, 1994; Kunkel, *et al*., 1991). The combination of these factors means that the possibility exists that the population of mtDNA molecules found within an individual can be diverse with the variants segregating and replicating independently.

It would seem then, that if all the individual mtDNA variants found in a mother were passed on to her offspring, then over time, the concept of an mtDNA "type," such as is seen today, would not exist. However, from the data accumulated via population genetic studies and forensic analysis, it is known that individuals usually have a single mtDNA type that can be distinguished from someone of another maternal lineage. Therefore, mechanisms, such as genetic bottlenecking must exist that restrict the level of mtDNA variation passed down to subsequent generations (Holland, 1999). Heteroplasmy has been estimated to occur at a relatively low rate, about 2-8% of the population as revealed by DNA sequence analysis (Holland, 1999). Probable reasons that mtDNA variation is now being observed at all is with the increased use of mtDNA analysis in forensic testing and the concordant development of databases. This was made possible by the advent of easier and faster automated sequencing methodologies. PCR and auto mated DNA sequencing has made it feasible for laboratories to apply large scale use of mtDNA analysis for forensic and data base use. Thus, more samples can now be studied with research emphasis placed on discordant results such as mutations when identity is an issue in the crimminal justice system.

Besides DNA sequencing other techniques have been employed in the study of human mtDNA mutations. These include Denaturing High Performance Liquid (DHPLC) (Underhill, *et al*., 1997) Denaturing Gradient Gel Electrophoresis (DGGE) (Steighner, *et al*., 1998; Hamekamp *et al*., 1996; Trulzch, *et al*., 1999; Tully, 1998).

Single S` tranded Conformational Polymorphism (Tully, 1998), Sequence Specific Oligonucleotides (Reynolds, 1999) and Temporal Gradient Gel Electrophoresis (Chen*, et al*., 1999). One study involving the use of DGGE, found heteroplasmy in 35 out of 253 individuals in the HVI region. Two of the subjects had detectable heteroplasmies at two different base pair positions (Tully, 1998). Mixing experiments performed in this study revealed that DGGE detects a mixture even when the minor contributor is 1% of the total. Direct sequencing can detect mixtures at a level of 10-15% minor contributors (Wilson*, et al*., 1997).

How does this then affect the interpretation of a mtDNA identity test? First, it must be understood that there are two categories of identity tests. The first one is matching the mtDNA type recovered from a piece of evidence (i.e. foreign hair found on a crime victim's body) to a known reference sample. In this test, referred to as an intracomparison analysis, a question is raised as to the identity of the mtDNA type associated with the crime scene. Nucleotide substitution should not effect this analysis because these substitutions are intergenerational. They occur during embryogenesis. Although once an individual has an mtDNA type, it is typically uniform in that person's genetic make-up, heteroplasmy can be a problem. One study documented the existence of differences in the mtDNA isolated from blood samples, buccal cell swabs, and hairs taken from the same individual (Wilson, *et al*., 1997). This study examined a family that

exhibited heteroplasmy at base pair position 16,.355 in the HVI region. One individual of that family revealed a changing pattern of heteroplasmy going from a predominantly thymine to a predominantly cytosine in one hair, when adjoining 2 cm. sections going from hair root to tip of that hair were examined. This changing pattern in heteroplasmy was not observed in the blood and buccal cell swabbings. Other members of that family exhibited different levels of heteroplasmy in different hair roots and shafts. The fact that blood and buccal cells have a multiple cellular origin while hair does not. These results also suggest that each mitochondrion is homoplasmic, but a cell might have at least two major mtDNA types` ` ` opposed to another school of thought that individual mitochondria are heteroplasmic` The Wilson paper (Wilson *et.al*., 1997) also illustrates the need for a forensic analyst to examine more than one reference sample per person (example, blood , buccal cell swabbings, and several hairs as reference material from any one individual).

Interestingly, heteroplasmy may actually help to strengthen a match in an intra–generational comparison . Because heteroplasmy is unusual, finding the same heteroplasmy between an unknown and a reference sample provides a good indication that both samples shared the same origin.

As opposed to the above example, where a comparison is performed to see if two items originated from the same source, testing is also performed to determine if two samples came from the same lineage. This type of analysis is most often used when identifying the physical remains of a missing person. When the remains of an unknown individual are recovered and a potential identity is assumed, reference samples from the suspected maternal relatives are collected, and a mtDNA comparison is performed

between the results obtained from the unknown remains and those obtained from the possible maternal relatives. If the two sequences are identical, then a match is called. But what if they differ by one base pair, or a heteroplasmy is observed but does not occur in either the remains or the reference samples? Interpretation becomes troublesome. Most forensic laboratories conducting mtDNA typing have particular criteria for calling a match. (AFIP Manual, FBI Manual). Typically, laboratories will not exclude on a one base pair difference. Rather they will call the result inconclusive, but cannot exclude. Only if the sequences differ at two or three bases will an exclusion be called.

The FBI mtDNA lab, for example, uses a two base or more exclusion rule. On the other hand, the AFIP DNA Identity lab has a three base exclusion rule.

Regarding the dispute between the FBI and the AFIP match criteria (2 base vs. 3 base differences), a small study was performed using the date accumulated in the project in order to help determine deviations where from Anderson and from each other in the 5 families studied. The results are as follows:

The deviations from Anderson range from 2 bases to 9 bases. This gives an average of 5.2 bases with a median of 4 bases. The national average as repeated by the FBI is six bases.

The base differences between the different families are broken down by family. These results are as follows:

PLI family

Mean base differences- 7 bases

Range- low-3 bases- high 10 bases

Median- 7.5 bases

Standard Deviation- 3

RGI family

Mean base difference- 10,5 bases

Range- low 4 bases- high 17 bases

Median- 10,5 bases

Standard Deviation- 4

RGII family

Mean base difference- 6.75 bases

Range- low 4 bases- high 13 bases

Median- 5 bases

Standard Deviation- 4

SMI family

Mean base difference- 7 bases

Range- low 3 bases- high 11 bases

Median – 7 bases

Standard Deviation- 4

GBI family

Mean base difference- 12.25 bases

Range- low 9 bases- high 17 bases

Median- 11.5 bases

Standard Deviation- 3.5

The sequence obtained from one family (SM1) was found to occur in 1.2 percent of the national mtDNA database (the database contains HV1 and HV2 sequences from 2426 individuals). All of the rest of the families, as well as the mutation sequence, were single source profiles.

It appears, from these results, that either the FBI or AFIP guidelines for not calling full exclusions are both very conservative and well thought out.

Until more is known about mitochondrial mutations, and due to the overall conservative nature of forensic biology, I made a decision was made in my laboratory to err on the conservative side and utilize the AFIP match criteria for full exclusions.

The sequences were then put into the national mtDNA database in order to ascertain the approximate frequency of each of the five family sequences and the

mutation. Currently, the database contains sequences from the HV1 and HV2 regions of 2426 unrelated individuals. Due to the non-Mendelian nature of mtDNA inheritance, a direct counting method is used in order to obtain some information on the frequency of any particular mtDNA sequence. The program that performs this is called "Mitosearch" and is available through the FBI or AFIP laboratories (Budowle*, et al*., in press).

To summarize the results all of the sequences analyzed ,except for one, were single source sequences. They were not found in the database of 2426 individuals. This includes the sequence with the mutation. One family sequence, the SMI family, was found to occur in 1.2% of the database. See **Addendix B** for more details.

From these results it appears that the interpretational guidelines as put forth by either the FBI or AFIP are both conservative in nature. However, because this research did reveal that two of the unrelated families (PLI and SMI) had a low of 3 base difference (see above) it was decided that the AFIP criteria was the more conservative and that was the one that I recommend be used.

CHAPTER III

CONCLUSIONS

The use of mtDNA analysis in human identification, though specialized, is now being used more often. In the United States, mtDNA testing is currently being performed by two public laboratories (FBI and AFIP) and five private laboratories (Biosynthesis, Inc., Bode Technologies, Lab-Corp., Mitotype, and Reliagene). The first U.S. court case where mtDNA evidence was introduced occurred in the state of Tennessee in 1996 (personal communication , Stewart, 1999). Since that time, mtDNA evidence has been admitted in court cases in at least ten states (Holland, *et al*., 1999). Outside the U.S., 40 other labs in eleven European countries are performing mtDNA testing. Great Britain has been conducting mtDNA testing since 1992 (Holland, *et al*., 1999). Thus, mtDNA analysis is now an accepted method of human identity testing worldwide. However, much is still to be learned. The major emphasis of this research was to gain a better understanding of the frequency of inter- maternal lineage DNA sequence changes and how they effect the interpretation of a forensic identity test. When this project was first designed the use of pedigree analysis was a novel approach. Since then, other researchers have used that method a have come up with similar results as this study. Other then its importance for forensic analysis these results might also have an impact on evolutionary studies as w The major emphasis of this research was to gain a better understanding of the frequency of inter - apparent disparity between thmaternal lineage phylogenetic mutation rate DNA sequence changes and how they effect th e interpretation of a for resolved. If the phylogenetic rates are accurate, then the common ancestor to the human

mitochondrial genome existed approximately 150,000 years ago (Horai, *et al*., 1999). However, if the pedigree analysis rates are accurate (which are, on the average, 200-fold higher than the phylogenetic rates) human mitochondrial ancestry began 750 years ago. Obviously, this cannot be true. What then causes this disparity? One explanation involves the fact that the oxidation reduction reactions occurring in the mitochondria produce free radicals that damage the DNA and the mitochondrial genome (Wallace,1999). Another explanation is that nuclear or mitochondrial gene products involved in mtDNA replication repair might have alleles that produce mutations (Larson & Clayton 1995). One paper (Suomalaninen, *et al*., 1995) found an autosomal gene mapped to chromosome 10Q that caused mitochondrial deletions. Still another explanation is that pedigree analysis detects mutations before they are eliminated by selection and not fixed (Paabo, 1996). Other explanations include the possibility is that nucleotide positions differ greatly in their tendency to mutate (Horai, *et al*., 1990). Most likely, some or all of these factors play a role.

Perhaps neither the phylogenetic rate nor the pedigree analysis rate is either right or wrong. Both may tell us different information on the same subject. Each could be used in different circumstances. For comparisons involving recent ancestry, like forensic analysis, or evolutionary/anthropological studies dating back to a common ancestor only hundreds of years ago, then the pedigree rate might be more applicable. For comparisons dating back hundreds of thousands or millions of years ago, a phylogenetic rate might be preferred.

A better understanding of the molecular biology of mtDNA repair and replication is needed. By developing purified enzymes used in mtDNA repair and replication, in

vitro assays could be developed to study mutational hot spots in mtDNA. This could lead to an elucidation of the biochemical basis for population variation and to a better understanding of the contribution that mutation and selection of mutations makes to genomic changes in mtDNA. One author has referred to the human mitochondrial genome as a microcosm of the human genome (Paabo, 1996). While the human mitochondrial genome was sequenced in 1981, eighteen years later scientists are still making new discoveries about it. This foreshadows what we can expect when the Human Genome Project is completed. Thus, the sequencing of the human mitochondrial genome was the beginning, not the ending, of a branch of study. It opened the door to whole new areas of study and analysis and to real world applications in the diverse fields of molecular biology, human evolution, medicine, and forensic science. This study was meant to contribute at least some knowledge to the field of mtDNA research with particular regard to its application in human genetic identity testing. But again, this study is only the beginning of what promises to be an exciting avenue of inquiry.

These studies dealt with the use of mtDNA sequencing analysis for forensic identity testing. The first study on the effects of environmental insults see **Appendix A** is typical of the kind of study needed to validate a technique before it can be implemented for laboratory use. This study revealed that mtDNA is very stable and not easily degraded. The PCR technique as well has proven to be a very robust method for analyzing mtDNA. The substitution rate study has added to the knowledge obtained from previous studies and observations. It has demonstrated the dynamic and still not fully understood nature of the mitochondrial genome. The substitution that was discovered (position 282-1 HVII) was a novel one; it had not been observed before. The mutation

rate (1 in 35 generations) corresponds extremely well to the two previous similar studies (1 in 33 and 1 in 40 generations). The similarity of the results in these studies indiciate that the intergenerational substitution rate of the HVI and HV II regions can now be better calculated and the results used to aid in the interpretation of a mtDNA based forensic test.

Appendix A

Environmental Insult Studies

DNA Typing of biological evidence collected at crime scenes has been greatly facilitated by PCR. The HVI and HVII regions of the human mitochondrial genome are two of the many target sites that are now available for examination because of the PCR technique. However, PCR does have its limitations. For example, heme from bloodstains, certain dyes from clothing, and humic acid from soil have all been shown to inhibit the reaction. (Akane, 1994). Furthermore, evidentiary material collected from a crime scene can be subjected to an unpredictable variety of external influences prior to its examination. Therefore, an understanding as to how some of these environmental insults can effect the ability of PCR to correctly amplify these samples is necessary.

Many previous studies have been done on genomic DNA markers (McNally, 1989a, McNally, 1989b; Cosso, 1995; Wallin, *et al*., 1998). Prior to 1995, when this research was performed, there had been no published data on the effects that environmental insults might have on the PCR application of human mtDNA. In 1995, the Technical Working Group of DNA Methods set up standard validations for DNA base markers used in human identity testing. Section 1.5.6 states:

> "Environmental studies evaluate the method using known or previously characterized samples exposed to a variety of environmental conditions. The samples should be selected to represent the types of specimens to be routinely analyzed by these methods. They should resemble actual evidence materials as closely as possible so that the effects of factors such as matrix, age, and degradative environment (temperature, humidity, UV) on a sample are considered."

The study herein described was the first step in addressing this issue for human mtDNA. This study was set up to develop an understanding of the effects that a variety

of variables would have on the amplification of the HVI and HVII regions of the D-loop of the human mitochondrial genome.

Most of the work involving the forensic use of human mtDNA has been done on skeletal remains. The osteoclast cells are in a very protected environment. Many of the usual environmental factors that might affect the integrity of a DNA sample, such as sunlight, soil, etc., would not have much direct effect on the DNA obtained from bone. The decision was made to use blood, not bone, teeth or hair, as the tissue from which the mt. DNA would be studied thereby allowing for a more direct effect of the environmental variables than would be seen on the protected bone DNA.

One of the aspects that sets apart the use of mt. DNA in forensic testing from genomic DNA is that there is no reliable way to quantify how much mt. DNA has been isolated from a sample. PCR technique has been shown to be extremely sensitive to the amount of DNA template needed for the reaction. (Hoevel, *et al*., 1992). For example, too much DNA can give erroneous results or totally inhibit the reaction. Too little DNA template will fail to produce an amplified product. Thus, another part of this study was to determine how much total human DNA was needed to produce the desired amplicon.

The results of this study have been presented once orally and once as a poster session at two separate meetings (Warren, 1996a, Warren, 1996b).

MATERIALS AND METHODS

Preparation of Samples and Contamination with a Variety of Chemicals

10 µl of blood was collected from a human donor. 50 µl of the liquid blood was then mixed with 50 µl of the following chemicals: isopropanol, soap solution, bleach, 70% ethanol, tap water, 0.85% saline, and an antiseptic spray. 50 µl of the blood/chemical mixture was then placed on a 100% non-bleached, sterile, white cotton cloth and allowed to dry for a period of 24 hours. Samples were made in duplicate and stored at -20° C.

Exposure to Environmental VariablesSunlight

Bloodstains were made up of 25 µl of liquid blood placed on sterile, non-bleached cotton cloth. The bloodstained cloth was placed in a Petri dish and exposed outdoors to sunlight for periods of one day, three days, three weeks, and five weeks. This was performed by placing the Petri dish containing the stained cloth onto the roof of a one story house. They were left there throughout the designated time frame for exposure. This was done in north central Texas in the month of April. The average daily temperature was 75-83 ° C. It rained once per week during the duration of the study. The weather was generally warm and sunny. The samples were removed from exposure at the appropriate time and stored at -20° C. All samples were made in duplicate.

Unprotected Exposure Outdoors

Bloodstains were prepared as described previously. They were placed outdoors in a Petri dish without a lid for periods of one, two, three, four, five, six, and seven days. The local they were placed in was a semi-wooded backyard in north central Texas. They were left outside throughout the duration of the study. The average daily day time

temperature was 75-83° C with night time temperatures of 45-60° C. During the period of exposure it rained for two days starting with the second day. The samples were removed at the appropriate time and stored at -20° C. All samples were made in duplicate.

Exposure to Heat

Bloodstains (see above) were placed in incubators with temperatures set respectively at 37° C and 56° C. The stains were kept in these incubators for periods of three days, one week, and three weeks in the 37° C incubator, and for one week, three weeks, and thirty-five days for the 56° C incubator. The samples were removed at the appropriate time period and stored at -20° C. Samples were made in duplicate.

Exposure to Humidity

Bloodstains (see above) were placed in a humidity chamber made by adding distilled water to blotting paper which was then placed into a sealed plastic container. The level of humidity was not monitored exactly but the levels were estimated to be between 50-90%. The paper was re-moistened on a daily basis. A strip of Parafilm™ was placed on top of the paper. The bloodstained cloth was then placed on the Parafilm™ in order to protect the cloth from coming into direct contact with the moist paper. One set of stains was kept at room temperature (approximately 25° C), and a second set was placed in a 37° C incubator. Samples were collected at one, three, and seven days and were stored at -20° C. All samples were made in duplicate.

Soil Exposure

Bloodstains (see above) were exposed to soil by burying them underground at a depth of approximately 2 feet. The location of the exposure was a backyard in north

central Texas. This was a semi-wooded location with trees, decaying vegetation and animal waste. The soil moisture was approximately 50%. Samples were collected at one, three, and seven days. The collected samples were stored at -20° C until processed. All samples were made in duplicate.

Exposure to Gasoline and Motor Oil

50 µl of liquid blood was mixed in ratios of 1 to 1 and 1 to 10 with commercial grade gasoline (87 octane) and 30W motor oil. 25 µl stains were made and placed on a 100% non-bleached, sterile, white cotton cloth. The stains were allowed to dry for a period of 24 hours. They were stored at -20° C until processed. All samples were made in duplicate.

Exposure to Ultraviolet (*uv*) Light

25 µl of whole blood was placed onto a piece of 100% non-bleached, sterile, white cotton cloth. The still-moist stain was placed under a germicidal *uv* light with a polychromatic wavelength of 280-320 nm. The stains were approximately 30 cm from the *uv* source. The stains were exposed for periods of one, five, ten, twenty, and fifty minutes. They were collected at the appropriate intervals and stored at -20° C. All stains were made in duplicate.

Non-Human Studies

Liquid blood from canine, feline, and equine sources were obtained from a veterinarian (personal communication, Warren, 1996). These samples were then tested with dog, cat and horse antisera to further check their authenticity. These samples were used to see if the human mtDNA primers would also give results on non-human samples that might be commonly encountered at a crime scene.

Sensitivity Studies

DNA commercially obtained (Life Technologies, Inc.) from the K562 human cell line was diluted into the following amounts: 40 ng/10μl, 20 ng/10 μl, 10ng/10 μl, 4 ng/ 10 μl, 2 ng/10 μl, 1 ng/10 μl, 0.4 ng/10 μl, 0.2 ng/10 μl, 0.1 ng/10 μl, and 0.04 ng/10 μl. The 0.04ng level was the limit of the detection method used and thus no value was seen in testing amounts of DNA less then this level. These were subjected to PCR amplification using human mt. DNA primers for the HVI region. This study was meant to determine the optimal amount of total cellular DNA template to be used in amplification reaction in order to get a PCR product.

Substrate Studies

25 μl of whole blood was placed onto a clean piece of denim cloth. The stain was allowed to dry for 24 hours and stored at -20° C. The butts of previously smoked, commercially available cigarettes were collected and stored at -20° C. The purpose of this study was to determine if some of the more typical substrata on which DNA is deposited at crime scenes would have any effect at all on the PCR amplification of HVI region of the human mitochondrial D-loop.

Extraction of DNA

The blood stains were cut into small pieces (approximately 0.5 to 1 cm. squares) and placed in a 400 μl of stain extraction buffer (10 mM Tris, 100 mM NaCL, 10 mM EDTA, 2% SDS, pH 8.0). 10 μl of Proteinase K (20mg/ml) was then added. (Life Technologies, Inc.). The solution was then placed in a 56° C heat block and incubated for a period not exceeding 24 hours. The fabric was then removed and 400 μl of Phenol/Chloroform/Isoamyl alcohol 25:24:1 (Life Technologies, Inc.) was used to extract

the DNA. The DNA was precipitated out of solution by adding 1 ml of cold 100% absolute alcohol and letting the solution sit at -20° C for 30 minutes. The sample was then centrifuged at 10,000 x g for 15 minutes. The alcohol was decanted off and the remaining alcohol removed by the use of a savant speed vacudryer for five minutes. The DNA was brought to solution by the addition of

36 μl of TE buffer, 10 mM Tris–HCL, 0.5 EDTA, pH 8.0 and placed in a 56° C heat block for two hours. The DNA sample was further concentrated and purified by treatment of the DNA with a Microcon 100™ filter column (Amicon, Inc.). 500 μl of TE buffer was added to the Microcon 100™ along with the DNA sample. It was then placed in a centrifuge for 15 minutes at 500 x g. The process was repeated. After the second centrifugation, the filter was inverted and centrifuged for two minutes at 1000 x g. The concentrated DNA is now ready for DNA quantitation.

DNA Quantitation

DNA quantitation was performed with a slot-blotting manifold (Life Technologies, Inc.) and chemiluminescence detection (ACES 2.0™ human DNA quantitation system, Life Technologies, Inc.).

The DNA was denatured by the addition of 100 μl of 0.5M NaOH and 0.5 μl of NaCl to 2μl of the sample. This was incubated for five minutes at room temperature. Simultaneously, a strip of Biodyne™ A nylon membrane (Life Technologies, Inc.) was soaked in 2X SSC for five minutes. The membrane was then placed in the slot-blot manifold system, and the samples were pipetted onto it. A vacuum was applied for five minutes. The membrane was removed and rinsed in 0.2M tris-HCL,pH7.5 and 2X SSC for a period of five minutes. The membrane was exposed to *uv* light for 90 seconds. A

human specific DNA probe (D17V1) was hybridized onto the membrane at 50° C. After hybridization and washes, Lumi-ples Plus™ (Life Technologies, Inc.) was applied to the membrane. The membrane was encased in a static-free plastic folder and placed in a film development cassette. It was incubated for approximately 18 hours. Kodak™ X-0MAT AR film was exposed to the membrane for 15 minutes and developed. The intensity of the reaction was compared to the intensity of the known DNA quantitation standards.

PCR Amplification

The primers used were specified to amplify a 287 base pair region of the HVI region of the D-loop. The base pair positions were 15971 to 16258. The sequence for the forward primer is as follows:

L strand - 5' TTA ACT CCA CCA TTA GCA CC 3'.

The reverse primer sequence is as follows:

H strand - 5' TGG CTT TGG AGT TGC AGT TG 3' (See table).

Amplifications were made up in a 50 µl reaction mix using a Techne 90Ô thermocycler. The reaction mix consisted of 10 pmol each of primer, 200 um each of dNTP, 10X reaction buffer (50mM KCL, 10mM Tris-HCl pH 8.3), 1.5mM $MgCl_2$, 0.001% gelatin), 1 unit of AmpliTaq™ DNA Polymerase, (PE Applied Biosystems), 4 µg of BSA and 2mg of DNA template).

The thermocycler conditions are as follows: 94° C for 30 seconds = 1 cycle; 94° C for 10 seconds, 62° C for 20 seconds, and 72° C for 10 seconds = 32 cycles; 4° C = soak.

Product Gel

In order to determine if the PCR reaction successfully amplified the DNA template, a product gel was made. The product gel consisted of 2X Agarose (Life

Technologies, Inc., DNA Typing Grade Agarose) and 1XTBE buffer (10X TBE Buffer, Life Technologies, Inc.). Using a mini-gel apparatus (Horizon 58 Horizontal Gel Electrophoresis Apparatus). The gel was run for 35 minutes at 150 V. Ethidium bromide was used to stain the gel. A 123 base pair ladder (Life Technologies, 1995) was used to determine the base pair migration of the PCR product. The PCR product would have to have migrated between the 250 and 300 base pair bands on the ladder.

RESULTS

Chemical Contaminants

With the exception of the bleach, treatment with all of the chemicals resulted in amplified DNA. Briefly, treatment with the 70% ethanol, soap solution, antiseptic spray, 25% saline, isopropanol, gasoline (both the 1:10 and 1:1 ratio), motor oil (both 1:10 and 1:1 ratio), failed to inhibit the PCR reaction.

Sunlight Treatment

The bloodstains that were left exposed to sunlight for three days, three weeks, and five weeks all produced PCR product. Interestingly, the only sample that did not produce PCR product was the sample that was exposed for just one day.

Exposure to Outdoor Environment

None of the samples left outdoors produced any amplified PCR product. It should be noted though, that neither was any DNA recovered during the DNA quantitation step. During the period of outdoor exposure, the samples were subjected to rain twice, and it is speculated that the rain diluted the bloodstains to the point beyond which the DNA could be recovered.

Exposure to Heat

The samples exposed to 37° C for three days, one week, and three weeks all produced properly amplified PCR product. The sample exposed to 50° C for seven days, 21 days, and 35 days also all produced properly amplified PCR product.

Exposure to Humidity

All of the bloodstains left in the humidity chamber at room temperature for the time periods of one ,three and seven days produced the properly amplified PCR product. The samples left in the humidity chamber at 37° C failed to produce any amplified DNA.

Exposure to Soil

No DNA was recovered from the samples exposed to soil. No amplification was observed when the extract solution was subjected to the PCR reaction. Thus, it appears that direct exposure to the soil of a biological fluid, such as blood, will indeed deteriorate the DNA to point beyond which even the PCR reaction can produce any recoverable amounts.

Exposure to UV Light

All of the samples exposed to UV light for the appropriate time (one, five ,ten, twenty, and fifty minutes) all produced amplified product.

Non-Human Studies

DNA isolated from the cat, dog, and horse bloodstains all failed to produce PCR product when amplified with primers that were said to be specific for the human mt. DNA HVI region.

Sensitivity Studies

The only sample that did not produce amplified PCR product was the sample that had 40 ng of DNA template added to it. This confirms empirical observations made by this researcher that too much DNA template added to PCR reaction could indeed have inhibitory effects. All of the other samples, including the 400 picogram amount, did produce amplified DNA products, thus also adding empirical evidence to the observation made that even small amounts of DNA can be amplified successfully.

Substrate Studies

Amplified product was obtained from both the denim and the cigarette butt samples. It is noted, however, that the denim sample required two attempts before amplification was successful. Once again, this observation is in accordance with other observations made by this researcher that denim is a difficult substrate to obtain a PCR product from.

Discussion

The results, in general, demonstrate the robustness of the PCR method to perform properly even when the DNA template has been treated in less than pristine conditions. This study also provides evidence that mtDNA can be a viable target for forensic DNA typing even when the sample has been exposed to conditions that are not likely to occur under general laboratory conditions.. Some interesting observations can be obtained from this study. DNA testing is still not entirely predictable. The results of the sunlight exposure study show that success is still sample- specific in that the only sample that did not amplify was the one exposed for the shortest period of time to sunlight.

The failure of the 37° C humidity treated samples confirms what many forensic biologists already know, that heat and humidity are the two main environmental variables that need to be controlled during the handling and storage of evidence.

Why the sample exposed to bleach did not amplify is not known. Speculations are that it could be due to the bleach's total denaturation of the DNA or to the fact that the impurities of the DNA could not be fully removed either by the ethanol precipitation or Mitrocon 100Ô treatment. It is noted that the extracted DNA always maintained a dark-brown tint indicating that the bleach/blood combination created impurities that co-precipitated in the DNA and inhibited the PCR.

Finally, the sensitivity studies revealed that high amounts of DNA do indeed inhibit PCR. Based on the results of this study, it was determined that 2ng of total human DNA was the optimal amount of DNA template to add to the PCR reaction in order to obtain successful amplification of the HVI region in D-loop of human mt. DNA.

Subsequent studies were performed by the FBI (Wilson, *et al*., 1995). That study carried this research one step further in that they attempted to sequence the mt. DNA after it had been subjected to environmental insults. The results of the FBI study revealed that no sequence changes were observed despite treatment with a variety of environmental insults.

Appendix B

Mitochondrial DNA Search Results by the Federal Bureau of Investigation

Search Profile
Profile ID: PLI Family Mutation

Search Range(s)	Differences from Anderson Sequence
16024 - 16365 73 - 340	73 G 263 G 282.1 T

Overall Search Results

Number of Differences From Search Profile	Number	Frequency	Cumulative Number	Cumulative Frequency
0	0	0	0	0
1	0	0	0	0
2	04	0.0016	4	0.0016
3	63	0.026	67	0.0276
4	165	0.068	232	0.0956
5	243	0.1002	475	0.1958
>5	1951	0.8042	2426	1

Search Profile
Profile ID: RGII

Search Range(s)	Differences from Anderson Sequence
16024 - 16365 73 - 340	16182 C 16183 C 16189 C 230 G

Overall Search Results

Number of Differences From Search Profile	Number	Frequency	Cumulative Number	Cumulative Frequency
0	0	0	0	0
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0

4	3	0.0012	3	0.0012
5	9	0.0037	12	0.0049
>5	2414	0.9951	2426	1

Search Profile
Profile ID: PLI

Search Range(s)	**Differences from Anderson Sequence**
16024 - 16365 73 - 340	73 G 263 G

Overall Search Results

Number of Differences From Search Profile	**Number**	**Frequency**	**Cumulative Number**	**Cumulative Frequency**
0	0	0	0	0
1	4	0.0016	4	0.0016
2	63	0.026	67	0.0276
3	164	0.0676	231	0.0952
4	244	0.1006	475	0.1958
5	293	0.1208	768	0.3166
>5	1658	0.6834	2426	1

Search Profile
Profile ID: SMI

Search Range(s)	**Differences from Anderson Sequence**
16024 - 16365 73 - 340	263 G 309.1 C 315.1 C

Overall Search Results

Number of Differences From Search Profile	**Number**	**Frequency**	**Cumulative Number**	**Cumulative Frequency**
0	30	0.0124	30	0.0124
1	134	0.0552	164	0.0676
2	202	0.0833	366	0.1509
3	183	0.0754	549	0.2263

4	147	0.0606	696	0.2869
5	208	0.0857	904	0.3726
>5	1522	0.6274	2426	1

Search Profile
Profile ID: RGI

Search Range(s)	**Differences from Anderson Sequence**
16024 - 16365 73 - 340	16182 C 16183 C 16189 C 16217 C 116 G 230 G 280 T 281 T

Overall Search Results

Number of Differences From Search Profile	**Number**	**Frequency**	**Cumulative Number**	**Cumulative Frequency**
0	0	0	0	0
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0
5	0	0	0	0
>5	2426	1	2426	1

Search Profile
Profile ID: GBI

Search Range(s)	**Differences from Anderson Sequence**
16024 - 16365 73 - 340	16129 A 16192 T 16270 T 16304 C 150 T 228 A 263 G 310 C 311 T

Overall Search Results

Number of Differences From Search Profile	Number	Frequency	Cumulative Number	Cumulative Frequency
0	0	0	0	0
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0
5	0	0	0	0
>5	2426	1	2426	1

REFERENCES

1. ABI Prism 310 Genetic Analyzer Users Manual 1998. PE Applied Biosystems, Perkin-Elmer Corp., Foster City, CA. Patent No. 903565, Rev. B.
2. ABI Prism Big Dye Terminator Cycle Sequencing Ready Reaction Kit. 1998. PE Applied Biosystems, Perkin-Elmer Corp., Foster City, CA.
3. ABI Prism DNA Sequencing Analysis Software Users Manual, PE Applied Biosystems, Perkin-Elmer Corp., Foster City, CA. Pat. No. 904532.
4. ABI Prism DNA Sequencing Chemistry Guide. 1995. Perkin-Elmer Corp., Foster City, CA. Pat. No. 903563, Version A.
5. ABI Prism Mitochondrial DNA Sequencing, Perkin-Elmer Corp., Foster City, CA.
6. Adachi, J., and M. Hasegawa. 1996. Tempo and mode of synonymous substitutions in the mitochondrial DNA of primates. *Mol. Biol. Evol.* 13: 200-208.
7. Adams, D. E, et al. 1991. Deoxyribonucleic acid (DNA) analysis by restriction fragment length polymorphisms of blood and other body fluid stains subjected to contamination and environmental insults. *J. For. Sci.* 36: 1284-1298.
8. Akane, A, et al. 1994. Identification of the heme compound copurified with deoxyribonucleic acid (DNA) from bloodstains, a major inhibitor of polymerase chain reaction (PCR) amplification. *J. For. Sci*. 39: 362-377
9. Alpney, L. 1997. *DNA Sequencing: From Experimental Methods to BioInformatics.* New York: Scientific Publishers, LTD – Springer-Verlag.

10. Anderson, S., et al. 1981. Sequence and organization of the human mitochondrial genome. *Nature* 209: 457-465.

11. Ankel-Simmon, F. and J. M. Cummings. 1996. Misconception about mitochondrial and mammalian fertilization: implications for theories on human evolution. *Proc. Natl. Acad. Sci. U.S.A.* 93: 13859-13863.

12. Armed Forces DNA Identification Laboratory. 1995. *Analysis of Mitochondrial DNA Sequencing Data*. Washington, D.C.: Office of the Armed Forces Medical Examiner.

13. Avise,JC. 1994 *Molecular Markers, Natural History and Evolution*. Chapman and Hall

14. Awadalla, P, et al. 1999. Linkage disequilibrium and recombination in hominid mitochondrial DNA. *Science* 286: 2524-2525

15. Ayala, F. 1995. The myth of Eve: molecular biology and human origins. *Science* 270: 1930-1936.

16. Bailliet, G, et al. 1994. Founder mitochondrial haplotypes in Amerindian populations. *Am. J. of Hum. Genet*. 55:27-33

17. Barrell, B. G., et al. 1979. A different genetic code in human mitochondria. *Nature* 282: 189-194.

18. Barrell, B. G., et al. 1980. Different pattern of codon recognition by mammalian mitochondria tRNAs. *Proc. Natl. Acad. Sci. U.S.A.* 77: 3167-3170.

19. Bendall, K.E. , et al. 1997. Heteroplasmic point mutations in the human mtDNA control region. *Am J Hum Genet* 59: 1276-1287

20. Bidocki, S. K., et al. 1997. Intracellular mitochondrial triplasmy in a patient with two heteroplasmic base changes. *Am J Hum Genet* 61: 1430-1438.

21. Biosynthesis, Inc. 1999. *Mitochondrial DNA Sequencing*. Lewisville, TX: Biosynthesis, Inc.

22. Bonatto, S.L. , Salanzo, F.M. , 1997. A single and early migration for the peopling of the Americas supported by mitochondrial DNA sequence data. *Proc Natl Acad Sci USA* 94: 1866-1871

23. Brown, M. D. 1990. Sequencing with Taq DNA polymerase. In *PCR Protocols: A Guide to Methods and Application*, Academic Press, 189-196.

24. Brown, M. D., and D. C. Wallace. 1994. Molecular basis of mitochondrial DNA disease. *J. Bioenergetics Biomembranes* 26: 273.

25. Brown, W. M. 1983. Evolution of animal mitochondrial DNA. In *Evolution of Genes and Proteins*, ed. M. Nei and R. Koehn Sinauier Press, 62-88.

26. Brown, W. M., et al. 1982. Mitochondiral DNA sequences of primates: tempo and mode of evolution. *J. Mol. Evol.* 18: 225-239.

27. Budowle, B., et al. 1990. Mitochondrial DNA: a possible genetic material suitable for forensic analysis. In *DNA and other Polymorphisms in Forensic Science*, ed. Henry C. Lee and R. E. Gaenssen, Year Book Medical Publishers, 76-97

28. Budowle, B. et al. In press. Mitochondrial DNA regions HVI and HVII population data. *Forensic Sci. Int*.

29. Butler, J. M. 1998. The use of capillary electrophoresis in genotyping STR loci. In *Forensic DNA Profiling Protocols*, ed. P. J. Lincoln, and J. Thomson,Humana Press, Totowa, NJ, 279-290.

30. Cann, R. L., et al. 1987. Mitochondrial DNA and human evolution. *Nature* 325: 31-36.

31. Case, J. T. and D. C. Wallace. 1981. Maternal inheritance of mitochondrial DNA polymorphisms in cultured human fibroblast. *Somatic Cell Genetics* 7: 103-108.

32. Chakraborty, R., and K. M. Weiss. 1991. Genetic variation of the mitochondrial DNA genome in American Indians is at mutation-drift equilibrium. *Amer. J. of Phys. Anthro.* 86: 497-506.

33. Chen, F. J., et al. 1999. Detection of mitochondrial DNA mutations by temporal temperature gradient gel electrophoresis. *Clinical Chem.* 45: 1162-1167.

34. Chen, Y., et al. 1995. Analysis of mtDNA variation in African populations reveals the most ancient of all human continent specific haplogroup. *Am. J. Hum. Genet.* 67: 133-149.

35. Clayton, D. A. 1982. Replication of animal mitochondrial DNA. *Cell* 28: 693-705.

36. Clayton, D. A. 1983. Transcription of the mammalian mitochondrial genome. *Ann. Rev. of Biochem.* 53: 573-594.

37. Comas, D., et al. 1995. Heteroplasmy in the control region of human mitochondrial DNA. *Genome Research* 5: 89-90

38. Connor, A. and Stoneking, M. 1994. Assessing ethnicity from human mitochondrial DNA types determined by hybridization with sequence specific oligonucleotides. *J. For. Sci*. 39: 1360-1371

39. Cosso, S. and R. Reynolds. 1995. Validation of the AmpliFLP DIS80 PCR amplification kit for forensic casework analysis according to TWGDAM guidelines. *J. For. Sci.* 40: 424-434.

40. Crouau-Roy, B., et al. 1996. A line scale comparison of the human and chimpanzee genomes: linkage, linkage disequilibrium and sequence analysis. *Hum. Mol. Genet*. 5: 1131-1137

41. Danan, C., D. Sternberg, et al. 1999. Evaluation of parental mitochondrial inheritance in neonates born after intracytoplasmic sperm injection. *Am. J. Hum. Genet.* 65: 463-473.

42. DiRenzo, A., and A. C. Wilson. 1991. Branching patterns in the evolutionary tree for human mitochondrial DNA. *Proc. Natl. Acad. Sci. U.S.A.* 88: 1597-1600.

43. Edwards, A., et al. 1992. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* 12: 241-253.

44. Edwards, A., et al. 1991. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am. J. Hum. Genet.* 49: 746-756.

45. Fagan, B.M. , 1990. *The Journey from Eden: The Peopling of Our World*. Thames and Hudson

46. FBI Laboratory DNA Unit II Mitochondrial DNA Sequencing Protocol. Federal Bureau of Investigation, Washington, DC 1997

47. Fisher, R. P., et al. 1992. DNA wrapping and bending by a mitochondrial high mobility group like transcriptional activator protein. *J. Biol. Chem.* 267: 3358-3367.

48. Forster, P., et al. 1996. Origin and evolution of the native American mtDNA variation: a reappraisal. *Am J Hum Genet*. 59: 935-945

49. Giles, R. E., et al. 1980. Maternal inheritance of human mitochondrial DNA. *Proc. of the Natl. Acad. Sci. U.S.A.* 77: 6115-6119.

50. Gill, P., et al. 1994. Identification of the remains of the Romanov family by DNA analysis. *Nature Genetics* 6: 130-135.

51. Gray, M. W. 1989. The evolutionary origins of organelles. *Trends in Genetics* 5: 294-299.

52. Gray, M. W., et al. 1999. Mitochondrial evolution. *Science* 283: 1476-1481.

53. Greenberg, J.H., et al. 1986. The settlement of the Americas: a comparison of the linguistic, dental and genetic evidence. *Curr Anthropol* 27: 477-479

54. Hammer, M.F. 1995. A recent common ancestry for human Y chromosomes. *Nature* 378: 376-378

55. Hanekamp, J.S., et al. 1996. Screening for human mitochondrial DNA polymorphisms with denaturing gradient gel electrophoresis. *Hum Genet* 98: 243-248

56. Harpending, H.C., et al. 1993. The genetic structure of ancient human populations. *Curr. Anthropol*. 34: 483-496

57. Hasegawa, M., and S. Horai. 1991. Time of the deepest root for polymorphism in human mitochondrial DNA. *J. Mol. Evol.* 32: 37-42.

58. Hasegawa, M., et al. 1993. Towards a more accurate estimate for the human mitochondrial DNA tree. *J. Mol. Evol.* 37: 347-354.

59. Hayashi, J. I., et al. 1985. Absence of extensive recombination between intra and inter species mitochondrial DNA in mammalian cells. *Experimental Cell Research* 160: 387-395.

60. Hayswirth, W. W., and P. J. Laipis. 1982. Mitochondrial DNA polymorphism in a maternal lineage of Holstein cows. *Proc. Natl. Acad. Sci. U.S.A.* 79: 4686-4690.

61. Higuchi, R., et al. 1984. DNA sequences from the Quagge, an extinct member of the horse family. *Nature* 312: 282-284.

62. Higuchi, R., et al. 1988. DNA typing from single hairs. *Nature* 332: 543-546.

63. Hoelzel, A. R., and A. Green. 1992. Population-level variation by sequencing PCR-amplified DNA. *Molecular Genetic Analysis of Population- A Practical Approach*. Ed. A. R. Hoelzel, Oxford Univ. Press, New York.

64. Holland, M. M., and T. J. Parson. 1999. Mitochondrial DNA sequence analysis – validation and use for forensic casework. *Forensic Sci. Rev.* 11: 21-50.

65. Holland, M. M., et al. 1993. Mitochondrial DNA sequence analysis of human skeletal remains: identification of remains from the Vietnam war. *J. For. Sci.* 38: 542.

66. Holland, M. M., et al. 1995. Mitochondrial DNA sequence analysis of human remains. *Crime Lab. Digest* 22: 109-115.

67. Holland, M. M., et.al. 1994. Identification of human remains using mitochondrial DNA sequencing: potential mother – child mutational events. *Advances in Forensic Haemogenetics*. 5: 399-407

68. Hopgood, R.,et al. 1992. Strategies for automated sequencing of human mitochondrial DNA directly from PCR products. *Biotechniques* 13: 83-92.

69. Horai, S., and K. Haysaka. 1990. Intraspecific nucleotide sequence differences in the major non-coding region of human mitochondrial DNA. *Am. J. Hum. Genet.* 46: 828-842.

70. Horai, S., et al. 1992. Man's place in Hominoidea revealed by mitochondrial DNA genealogy. *J. Mol. Evol.* 35: 32-43.

71. Horai, S., et al. 1993. Peopling of the Americas founded by four major lineages of mitochondrial DNA. *Mol. Biol. Evol.* 10: 23-47.

72. Horai, S., et al. 1995. Recent African origin of modern humans revealed by complete sequences of Hominoid mitochondrial DNAs. *Proc. Natl. Acad. Sci. U.S.A.* 92: 532-536.

73. Horai,S. , et.al. 1991. Phylogenetic affiliation of ancient and contemporary humans inferred from mitochondrial DNA. *Phil Trans R Soc Lon Ser B* 333: 409-500

74. Howell, N. 1997. mtDNA recombination: what do In vitro data mean? *Am. J. Hum. Genetics* 61: 19-22.

75. Howell, N., and D. Mackey. 1997. Reply to Macauley, et al. *Am. J. Hum. Genet.* 61: 986-990.

76. Howell, N., et al. 1996. How rapidly does the human mitochondrial genome evolve? *Am. J. Hum. Genet.* 59: 501-509.

77. Ivanov, P. L., et al. 1996. Mitochondrial DNA sequence heteroplasmy in the Grand Duke of Russia Georgij Romanov establishes the authenticity of the remains of Tsar Nicholas II. *Nature Genetics* 12: 417-420.

78. Johnson, I. D. 1996. Introduction to fluorescence techniques. Molecular Probes Handbook.

79. Jorde, L.B., et al. 1995. Origins and affinities of modern humans: a comparison of mitochondrial and nuclear genetic data. *Am. J. of Hum. Genet*. 57:523-538

80. Jorde, L.B., et al. 1997. Microsatellite diversity and the demographic history of modern humans. *Proc. Natl. Acad. Sci.* USA 94:3100-3103

81. Jorde, L.B., et al. 1998. Using mitochondrial and nuclear markers to reconstruct human evolution. *BioEssays* 20: 126-136

82. Kirby, L. T. 1990. *DNA Fingerprinting: An Introduction*. New York: Stockton Press.

83. Kocher, T. D., et al. 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. U.S.A.* 86: 6196-6200.

84. Krings, M., et al. 1998. Neanderthal DNA sequences and the origin of modern humans. *Cell* 90: 19-30.

85. Kunkel, T. A., and L. A. Loeb. 1981. Fidelity of mammalian DNA polymerases. *Science* 213: 765.

86. Kurosaki, K., et al. 1993. Individual DNA identification from ancient human remains. *Am. J. Hum. Genet.* 53: 638-643.

87. Larson, N. and D. A. Clayton. 1995. Molecular genetic aspects of human mitochondrial disorders. *Ann. Rev. Genetics* 29: 151-178.

88. Levin, B. C., et al. 1999. A human mitochondrial DNA reference material for quality control in forensic identification, medical diagnosis and mutation detection. *Genomics* 55: 135-146.

89. Lundstrom, R., et al. 1992. Estimating substitution rates from molecular data using coalescent. *Proc Natl Acad Sci* 89: 5961-5965

90. Macaulay, V. A. 1997. Mitochondrial DNA mutation rate – no need to panic. *Am. J. Hum. Genet.* 61: 986-990.

91. Maddison, D. R. 1991. African origin of human mitochondrial DNA reexamined. *Syst. Zool.* 40: 355-363.

92. Maddison, D. R., et al. 1992. Geographic origins of human mitochondrial DNA: phylogenetic evidence from control region sequences. *Syst. Biol.* 41: 111-124.

93. Manfredi, G., et al. 1997. The fate of human sperm-derived mtDNA in somatic cells. *Am. J. of Hum. Gen.* 61: 953-960.

94. McNally, L., et al. 1989a. Evaluation of deoxyribonucleic acid (DNA) isolated from human bloodstains exposed to ultraviolet light, heat, humidity, and soil contamination. *J. For. Sci.* 34: 1059-1069.

95. McNally, L., et al. 1989b. The effects of environment and substrata on deoxyribonucleic acid (DNA): the use of casework samples from New York City. *J. For. Sci.* 34: 1070-1077.

96. Melnick, D. J., et al. 1992. Mitochondrial DNA: its uses in anthropological research. *Molecular Applications in Biological Anthropology*, ed. E. J. Devor (Cambridge Univ. Press), 179-233.

97. Melton, T. , et al. 1997a. Extent of heterogeneity in mitochondrial DNA of European populations. *J. For. Sc*i. 42: 582-592

98. Melton, T., and M. Stoneking. 1996. Extent of heterogeneity in mitochondrial DNA of ethnic Asian population. *J. For. Sci.* 41: 591-602.

99. Melton, T.,et al. 1997b. Extent heterogeneity in mitochondrial DNA of Sub-Saharan African populations. *J. For. Sc*i. 42;582-592

100. Merriwether, D. A. 1991. The structure of human mitochondrial DNA variation. *J. Mol. Evol.* 33: 543-555.

101. Merriwether, D.A., et al. 1995. Distribution of the four founding lineage haplotypes in Native Americans suggests a single wave of migration for the new world. *Am J Phys Anthropol* 98: 411-430

102. Michikawa, Y., et al. 1999. Aging- dependent large accumulation of point mutations in the human mtDNA control region for replication. *Science* 286: 774-779

103. Morin, P.A., et al. 1994. Kin selection, social structure, gene flow and the evolution of chimpanzees. *Science.* 265: 1193-1201

104. Mountain, J. L. 1995. Demographic history of India and mtDNA sequence diversity. *Am. J. Hum. Genet.* 56: 979-992.

105. National Research Council. 1996. *The Evaluation of Forensic DNA Evidence*. National Academy Press.

106. Nei, M., 1992. Age of the common ancestor of the human mitochondrial DNA. *Mol. Biol. Evol.* 9: 1176-1178.

107. Nei, M. and Livshits, G. 1989. Genetic relationships of Europeans, Asians and Africans and the origin of modern *Homo sapiens. Hum. Hered*. 39: 276-281

108. Newton, C. R., and A. Graham. 1997. *PCR 2nd Edition*. New York: Bios Scientific Publish, Inc. Springer-Verlag.

109. Paabo, S. 1989. Ancient DNA extraction, characterization, molecular cloning and enzymatic amplification. *Proc. Natl. Acad. Sci. U.S.A.* 86: 1939-1943.

110. Paabo, S. 1996. Mutational hot spots in the mitochondrial microcosm. *Am. J. Hum. Genet.* 59: 493-496.

111. Paabo, S., et al. 1988. Mitochondrial DNA sequences from a 7,000 year old brain. *Nucleic Acid Research* 16: 9775-9787.

112. Paabo, S., et al. 1989. Ancient DNA and the polymerase chain reaction. *J of Biol Chem* 264: 9709-9712.

113. Parson, T.J., et al. 1997. A high observed substitution rate in the human mitochondrial DNA control region. *Nature Genetics* 15: 363-368

114. Pellekaan, S.M., et al. 1998 Mitochondrial control region sequence variation in Aboriginal Australians. *Am J. Hum. Genet*. 62: 435-439

115. Penny, D., et al. 1995. Improved analysis of human mtDNA sequences support a recent African origin for *Homo sapiens*. *Mol. Biol. Evol.* 12: 863-882.

116. Pesole, G., et al. 1992. The evolution of the mitochondrial d-loop region and the origin of modern man." *Mol. Biol. Evol.* 9: 587-598.

117. Reynolds, R. 1999. Heteroplasmic mutations in the mtDNA control region as detected by SSO probes. Presented at the *10th Annual Symposium on Human Identification*. Orlando, Florida

118. Rickards, O., et al. 1999. Mitochondrial DNA history of the Cayapa Amerinds of Ecuador: detection of additional founding lineages for the Native American populations. *Am. J. Hum. Genet.* 65: 519-530.

119. Rogan, P. K., and J. J. Salvo. 1990. Molecular genetics of pre-Columbian South American mummies. *Molecular Evolution* 223-234

120. Rogan, P. K., and J. J. Salvo. 1990. Study of nucleic acids isolated from ancient human remains. *Yearbook of Physical Anthropology* 33: 195-214.

121. Rogers, A.R. and Jorde, L.B. 1995. Genetic evidence on the origins of modern humans. *Hum. Biol*. 67: 1-36

122. Ruvolo, M. 1996. A new approach to studying modern human origins: hypothesis testing with coalescence time distributions. *Mol. Phylogenet. Evol.* 5: 202-219.

123. Ruvolo, M., et al. 1993. Mitochondrial CoII sequences and modern human origins. *Mol. Biol. Evol.* 10: 1115-1135.

124. Saiki, R. K., et al. 1988. Primer-directed enzymatic amplification of DNA with thermostable DNA polymerase. *Science* 239: 487-491.

125. Sajantilla, A. 1995. Genes and languages in Europe: An analysis of mitochondrial lineages. *Genome Res.* 5: 42-52.

126. Sanger, F., et al. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* 74: 5467.

127. Shadel, G. S. and D. A. Clayton. 1993. Mitochondrial transcription, initiation, variation and conservation. *J. Biol. Chem.* 268: 1608-1686.

128. Steighner, R.J. and Holland M. 1998. Amplification and sequencing of mitochondrial DNA in forensic casework. *Methods in Mol Biol* 98: 213-218

129. Stone, A.C. , Stoneking, M. 1998. MtDNA analysis of a prehistoric Oneota population: implications for the peopling of the new world. *Am. J. of Hum Genet*. 62: 1153-1170

130. Stoneking, M., et al. 1990. Geographic variation in human mtDNA from Papua New Guinea. *Genetics* 124: 717-733.

131. Stoneking, M., et al. 1991. Population variation of human mtDNA control region sequences detected by enzymatic amplification and SSO Probes. *Am. J. Hum. Genet.* 48: 370-382.

132. Stoneking, M., et al. 1992. New approaches to dating suggest recent age for human mitochondrial DNA ancestor. *Phil. Trans. R. Soc. London* 337: 167-175.

133. Sullivan, K. M., et al. 1991. Autosomal amplification and sequencing of human mitochondrial DNA. *Electrophoresis* 12: 17-21.

134. Suomalainen, A., et al. 1995. An autosomal locus predisposing to deletions of mitochondrial DNA. *Nature Genetics* 9: 146-151

135. Szathmary, E.J.E. 1993. Genetics of aboriginal North Americans. *Evol Anthropol* 1:202-220

136. Takahata, N. 1993. Allelic genealogy and human evolution. *Mol. Biol. And Evol*. 10:2-22

137. Takahata, N., et al. 1995. Divergence time and population size in the lineage leading to modern humans. *Theor. Popul. Biol*. 28: 198-221

138. Tamura, K., and M. Nei. 1992 Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10: 512-526.

139. Templeton, A.R. 1993. The Eve hypothesis: a genetic critique and reanalysis. *Am Anthropol.* 95: 51-72

140. Thayagarajan, B., et al. 1996. Mammalian mitochondria possess homologous DNA recombination activity. *J. Biol. Chem.* 271: 27536-27543.

141. Torroni, A., and Wallace, D.C. 1995. Mitochondrial DNA haplogroups in Native Americans. *Am J Hum Genet* 56: 1234-1236

142. Torroni, A., et al. 1993. Asian affinities and continental radiation of the four founding Native American mtDNAs. *Am. J. Hum. Gen.* 53: 563-590.

143. Torroni, A., et al. 1994. Mitochondrial DNA clock for the Amerinds with its implications for timing their entry into North America. *Proc. Natl. Acad. Sci. U.S.A.* 91: 1503-1507.

144. Trulzsch, B., et al. 1999. DGGE is more sensitive for the detection of somatic point mutations than direct sequencing. *BioTechniques* 27: 266-268

145. Tully, L.A. 1998. *Examination of the Use of Forensic DNA Typing from Two Perspectives: I. Mitochondrial Heteroplasmy; II. The Role of DNA Typing in Criminal Investigations.*- Doctoral Dissertation University of Maryland, Baltimore, MD.

146. Underhill, P. A., et al. 1997. Detecting numerous Y chromosome allelic polymorphisms by denaturing high performance liquid chromatography. *Genomic Res.* 10: 996- 1001

147. Vigilant, L., et al. 1989. Mitochondrial DNA sequences in single hairs from a Southern African population. *Proc. Natl. Acad. Sci. U.S.A.* 86: 9350-9354.

148. Vigilant, L., et al. 1991. African populations and the evolution of human mitochondrial DNA. *Science* 233: 1503-1507.

149. Wakeley, J. 1993. Substitution rate variation among sites in the hypervariable region of human mitochondrial DNA. *J. Mol. Evol.* 37: 613-623.

150. Wallace, D. C. 1999. Mitochondrial diseases in man and mouse. *Science* 283: 1482-1488.

151. Wallace, D. W. 1989. Mitochondrial DNA mutations and neuromuscular disease. *Trends in Genetics* 5: 9-13.

152. Wallace, DC. and Torroni, A. 1992. American Indian prehistory as written in the mitochondrial DNA: A Review. *Hum Biol* 64: 403-416

153. Wallin, J. M., et al.1998. TWGDAM validation of the AMPFISTR Blue PCR amplification kit for forensic casework analysis. *J. For. Sci.* 43: 854-870.

154. Walsh, P.S., et al. 1991. Chelex 100™ as a medium for simple extraction of DNA for PCR- based typing from forensic material. *BioTechniques* 10: 506-513

155. Ward, R. H., et al. 1991. Extensive mitochondrial DNA divrsity in a single Amerindian tribe. *Proc. Natl. Acad. Sci. U.S.A.*88: 8720-8724.

156. Warren, J. E., et al 1996a. The effects of various types of environmental insults on the amplification of human mitochondrial DNA. *Proc. Am. Acad. Of Forensic Science*, Abstract B56. Nashville, TN.

157. Warren, J. E., et al. 1996b. The effects of various types of environmental insults on the amplification of human mitochondrial DNA. *Proc. from the 7th Inv. Symposium on Human Identification* 196. Scottsdale, Arizona

158. Willis, C. 1995. When did Eve live? An evolutionary detective story. *Evolution* 49: 593-607.

159. Wilson, M. R., et al. 1995a.Validation of mitochondiral DNA sequencing for forensic casework analysis. *Int. J. Legal Med.* 108: 68-74.

160. Wilson, M. R., et al. 1995b. Extraction, PCR amplification and sequencing of mitochondrial DNA from human hair shafts. *BioTechniques* 18: 662-669.

161. Wilson, M. R., et al. 1997. A family exhibiting heteroplasmy in the human mitochondrial DNA control region reveals both somatic mosaicism and pronounced segregation of mitotype. *Hum. Genet.* 100: 172-181.

162. Yang, Z. 1995. A space-time process model for the evolution of DNA sequences. *Genetics* 139: 993-1005

163. Zhang, N., and E. S. Young. 1996. Genetic typing by capillary electrophoresis with the allelic ladder as an absolute standard. *Anal. Chem.* 68: 2927-2931.

nucleotide substitutions/ site/ myr. This mutation rate was very consistent with several other similar studies. This increased mutation rate needs to be considered by forensic testing laboratories performing mtDNA sequence analysis prior to formulating any conclusive results.